AN
SSC
PUBLICATION

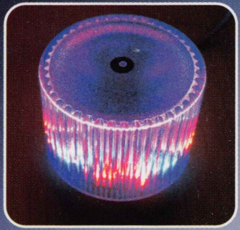
Stop Forged Spam with SPF

LINUX JOURNAL

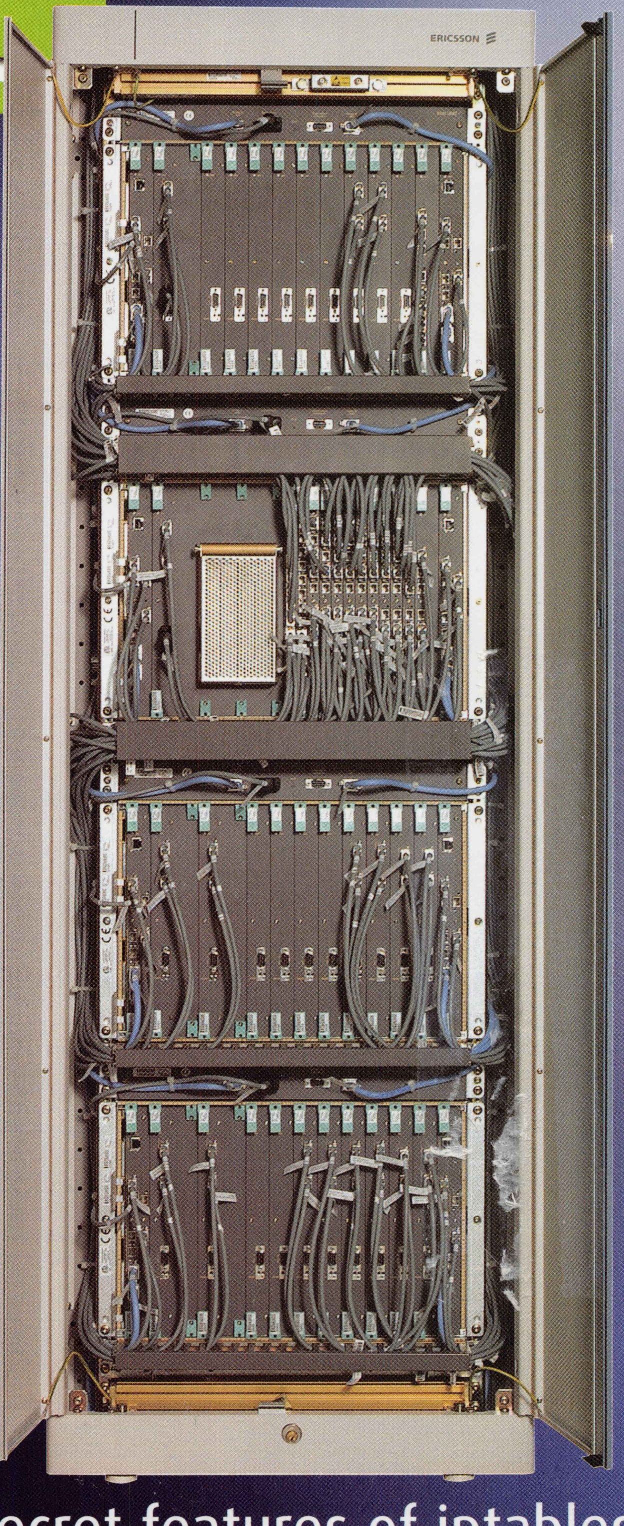
The Monthly Magazine of the Linux Community • APRIL 2004

IT'S ALIVE!

A SIMPLE USB DRIVER PROJECT



Distributed Security for Telecom Clusters



Samba logging for security audit trails

Top 8 PHP security bugs: protect your app

- > Put your server behind a Zorp proxy
- > Point-and-click mail encryption
- > Add a blog to your Web site: COREBlog
- > How IT tradeshows can get a clue

USA $5.00 CAN $6.50
www.linuxjournal.com
04
0 71486 03102 4

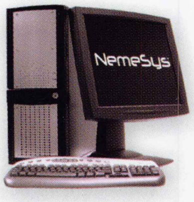
Tar pit and other secret features of iptables

Experience the most computing power in the smallest footprint possible...

AMD Opteron™

The RackSaver® BladeRack™, powered by AMD Opteron™ processors, is designed to support a wide variety of enterprise solutions...including yours.

Over the past 10 years, RackSaver has provided customers in every industry from Entertainment to Oil and Gas with the most powerful, reliable, and scalable high performance clusters available on the market today. Our customers understand that by increasing density in their server room they improve productivity while lowering their overall total cost of ownership. In addition, RackSaver offers complete onsite enterprise-level support 24 hours a day, 7 days a week to maximize your investment protection. Discover why industry leaders like Industrial Light and Magic, NVIDIA, Shell, WesternGeco, and Novartis rely on RackSaver for all of their high performance computing needs.



THE NEMESYS

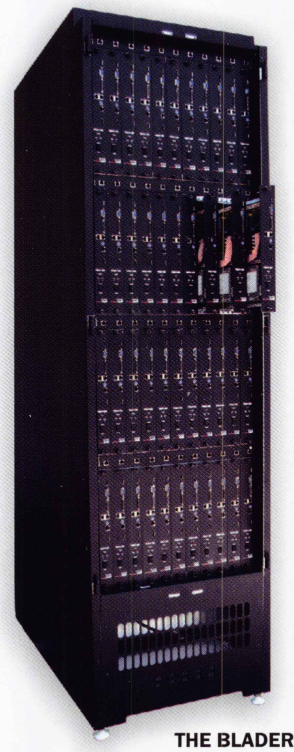
Did we mention that we sell high-end customizable graphics workstations too?

RackSaver has launched a new line of high-end graphic workstations. NemeSys™, powered by the NVIDIA Quadro® FX family, takes the leading design and digital content creation applications to a new level of interactivity. For the first time, styling and production rendering become integral functions of the design workflow, shortening the production process and enabling faster time to market.

THE HIVE CLUSTER

Need a blade solution that can fit under your desk at work and provide more power than ever before?

The HIVE™ houses up to 11 blades that can accommodate a variety of options including high-end servers, SAN, NAS-based storage, along with InfiniBand®. Both the 8 blade and 11 blade versions are fully compatible with RackSaver's RS-1100VM BladeRack line of products. The HIVE's innovative cooling technologies are specifically designed for the needs of small rack systems and also make them suitable for a wide range of environments including military, industrial or a simple office setting.

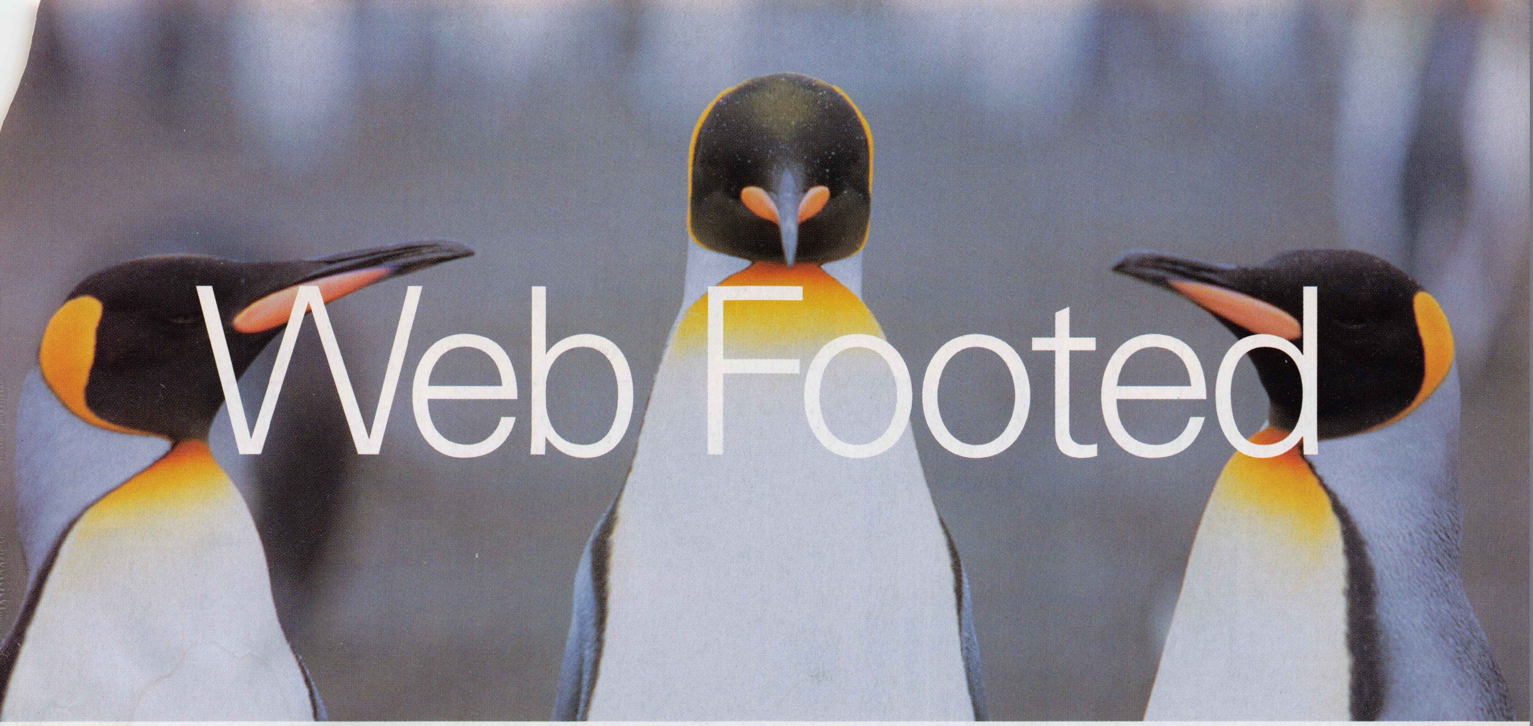


THE BLADERACK RS-1100VM

RackSaver®
Leader in High-Density Servers!

Start configuring your company's high performance supercomputing solution now by visiting RackSaver online at www.racksaver.com or call one of our Account Managers at (888) 942-3800.

Copyright ©2003, RackSaver Inc. All Rights Reserved. RackSaver, the RackSaver logo, BladeRack, and NemeSys are trademarks or registered trademarks of RackSaver Inc. Intel and Xeon and combinations thereof are trademarks of Intel Corporation or its subsidiaries in the United States or other countries.



The right management software allows you to seamlessly integrate Linux into your infrastructure.

Linux Management Software

No matter what type of Linux implementation you have planned, you'll still need management software to ensure that it's seamlessly integrated across your enterprise. Our management software solutions help maximize the potential of Linux across your entire business by letting you take full advantage of its unique scalability and open-source programming. As a result, you can increase productivity while decreasing total cost of ownership. To find out more about our continuing commitment to Linux innovations, go to ca.com/linux today.

ca Computer **Associates**®

© 2003 Computer Associates International, Inc. (CA). All rights reserved.

Is security this big when managing your data center?

Biometric Authentication
IP Filtering
Encryption (SSH)
Auditing

Ours is, only in smaller packages.

Fingerprint Authentication Scanner
AlterPath™ Bio

Enterprise KVM Solutions
AlterPath™ KVM

Advanced Console Servers
AlterPath™ ACS

Network Management Gateway
AlterPath™ Manager

Intelligent Power Distribution Units
AlterPath™ PM

Cyclades' data center management solutions offer a full range of security features across its entire product line of console servers, power management, KVM, biometric scanner and network management. With SSH v2, IP Filtering, strong authentication, event logging and data logging, Cyclades can make your network into a secure heavyweight contender in the data center world.

LINUX INSIDE

For a **FREE** white paper on data center security, please visit us at www.cyclades.com/securitywp

LINUX JOURNAL READERS' CHOICE 2003
"Best Hardware for Linux since 1995"

www.cyclades.com/lja

1.888.cyclades ▪ 1.888.292.5233 ▪ sales@cyclades.com

cyclades
Everywhere with Linux

©2004 Cyclades Corporation. All rights reserved. All other trademarks and product images are property of their respective owners. Product information subject to change without notice.

LINUX JOURNAL

APRIL 2004 ISSUE 120

COVER STORY

68 SECURITY DISTRIBUTION FOR LINUX CLUSTERS

One of our recurring themes over the past year has been how to enable new, finer-grained security models that go beyond the old-school users and groups. This month, take a step into the world of process-level access control on page 68.

FEATURES

54 REAL-WORLD PHP SECURITY

Learn the top four PHP security mistakes and the three key techniques you can use to secure your PHP app.

XAVIER SPRIET

62 SPF OVERVIEW

Spam, scams and worms all use e-mail forgery. Put a stop to it with the new mark of quality for your domain.

MENG WENG WONG

68 SECURITY DISTRIBUTION FOR LINUX CLUSTERS

Extend Linux Security Modules to enforce security rules across many systems.

IBRAHIM HADDAD AND MIROSLAW ZAKRZEWSKI

INDEPTH

82 CONSTRUCTING RED HAT ENTERPRISE LINUX V. 3

Behind the scenes, contentious IT firms have their say in a new high-end distribution.

TIM BURKE

88 SAMBA LOGGING FOR AUDIT TRAILS

When you have high-security audit requirements, use the source and add custom log entries.

EDWARD S. KABLAOUI

EMBEDDED

46 DRIVING ME NUTS

Writing a Simple USB Driver

GREG KROAH-HARTMAN

TOOLBOX

20 AT THE FORGE

Weblogs

REUVEN M. LERNER

24 KERNEL KORNER

The Hidden Treasures of iptables

CHRIS LOWTH

30 COOKING WITH LINUX

François, Can You Keep a Secret?

MARCEL GAGNÉ

36 PARANOID PENGUIN

Application Proxying with Zorp, Part II

MICK BAUER

COLUMNS

52 LINUX FOR SUITS

Showtime

DOC SEARLS

96 EOF

SOLIS, a Brazilian Free Software Cooperative

CESAR BROD

REVIEW

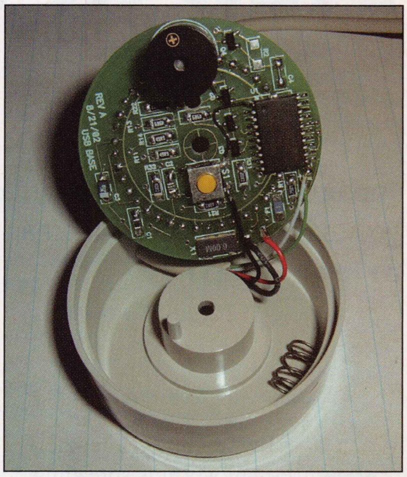
18 INTRUSION DETECTION WITH SNORT, APACHE, MYSQL, PHP, AND ACID

SEAN TIERNEY

DEPARTMENTS

4 FROM THE EDITOR
6 LETTERS
10 UPFRONT
14 FROM THE PUBLISHER
75 ON THE WEB
76 BEST OF TECHNICAL SUPPORT
80 NEW PRODUCTS
81 ADVERTISERS INDEX
95 MARKETPLACE

COVER IMAGE OF THE TELECOM SERVER PLATFORM (TSP) COURTESY OF ERICSSON COMMUNICATIONS INC.



Three wires, three LEDs—how hard can the device driver be? Page 46.

NEXT MONTH

SYSTEM ADMINISTRATION

Think you've got e-mail problems? Ludovic Marcotte took on the task of creating a new mail infrastructure for a university with 35,500 users and 125,000 incoming messages per day. And up to 80% of the incoming mail traffic is spam. OpenLDAP, Cyrus and Postfix to the rescue.

An unhappy fact about upgrades is that sometimes you have to back them out—fast. James Olin Oden has the next best thing to a time machine. Learn how to set up transactions in the RPM package manager so that you can roll them back.

Aaron Trumm has recorded and released seven albums, and has a new project—a Linux-based recording studio based on the free digital audio workstation program Ardour.

Security One Step at a Time

The attack path between intruders and your data might be shorter than you think.

BY DON MARTI

As I write this, yet another e-mail worm is spreading among non-Linux computers and incidentally filling my mailbox with "YOU HAVE A VIRUS" bounces from dumb software that somehow doesn't yet get the concept that worms forge mail. There's nothing like a worm attack that spares Linux to bring out the smug superiority in Linux users.

Cut it out. The attack path here is one step long. All that's keeping us safe is that most programs for Linux don't make it easy to run attachments from incoming mail. But combine the right vulnerability in a common desktop app with a little social engineering, and you've got a Linux worm.

Last year, the not-so-dramatically-named CAN-2003-0434 vulnerability allowed humble PDF files to run arbitrary commands as you. Linux users and distributions dealt with it quickly enough that it didn't turn into a vector for spreading a worm. With today's larger Linux user base and more desktop standardization, the next vulnerability will be a bigger risk.

Now that we've scared you, we'll cover the tools you could use to prevent not just a mail worm, but other attacks we don't know about yet. Run a local firewall and don't let programs on your company's desktops reach outside SMTP servers. Deploy exactly the firewall policy you want, on every host, with the advanced iptables advice in Chris Lowth's Kernel Korner on page 24. As you move your business apps to PHP, design them for security with Xavier Spriet's battle-tested designs on page 54.

And, make the next move in the spam wars. Deal with forgery where it starts. Although the US has essentially legalized spam, all the ISP advertising we've seen recently has used spam filtering as a selling point. Sender Permitted From, which Meng Weng Wong covers on page 62, lets you pop up out of the weeds and get mail through to customers who use strict spam filtering. SPF is a "look at me, I'm legit" measure you can deploy in a few minutes for a simple mail configuration.

Finally, in our cover story, Ibrahim Haddad and Miroslaw Zakrzewski explain a promising example of how to apply the kernel's Linux Security Module (LSM) interface to add process-level access control for telecom apps running on clusters (page 68). Developers can carry out this level of work, free of restrictions, because of the freedom that the GPL licensing consensus gives all of us. Keep your systems secure and enjoy this month's issue.

Don Marti is editor in chief of *Linux Journal*.

LINUX JOURNAL

APRIL 2004
ISSUE 120

EDITOR IN CHIEF Don Marti, ljeditor@ssc.com
MANAGING EDITOR Jill Franklin, jill@ssc.com
SENIOR EDITOR Doc Searls, doc@ssc.com
SENIOR EDITOR Heather Mead, heather@ssc.com
ART DIRECTOR Garrick Antikajian, garrick@ssc.com
TECHNICAL EDITOR Michael Baxter, mab@cruzio.com
SENIOR COLUMNIST Reuven Lerner, reuven@lerner.co.il
CHEF FRANÇAIS Marcel Gagné, mggagne@salmar.com
SECURITY EDITOR Mick Bauer, mick@visi.com
CONTRIBUTING EDITORS
David A. Bandel • Greg Kroah-Hartman • Ibrahim Haddad • Robert Love • Zack Brown • Dave Phillips • Marco Fioretti
PROOFREADER Geri Gale

VP OF SALES AND MARKETING Carlie Fairchild, carlie@ssc.com
MARKETING MANAGER Rebecca Cassity, rebecca@ssc.com
INTERNATIONAL MARKET ANALYST James Gray, jgray@ssc.com
REGIONAL ADVERTISING SALES
NORTHERN USA: Joseph Krack, +1 866-423-7722 (toll-free)
EASTERN USA: Martin Seto, +1 905-947-8846
SOUTHERN USA: Annie Tiemann, +1 866-965-6646 (toll-free)
ADVERTISING INQUIRIES ads@ssc.com

PUBLISHER Phil Hughes, phil@ssc.com

WEBMASTER Willy Smith, webmaster@ssc.com
TECHNICAL MANAGER Dan Wilder, dan@ssc.com

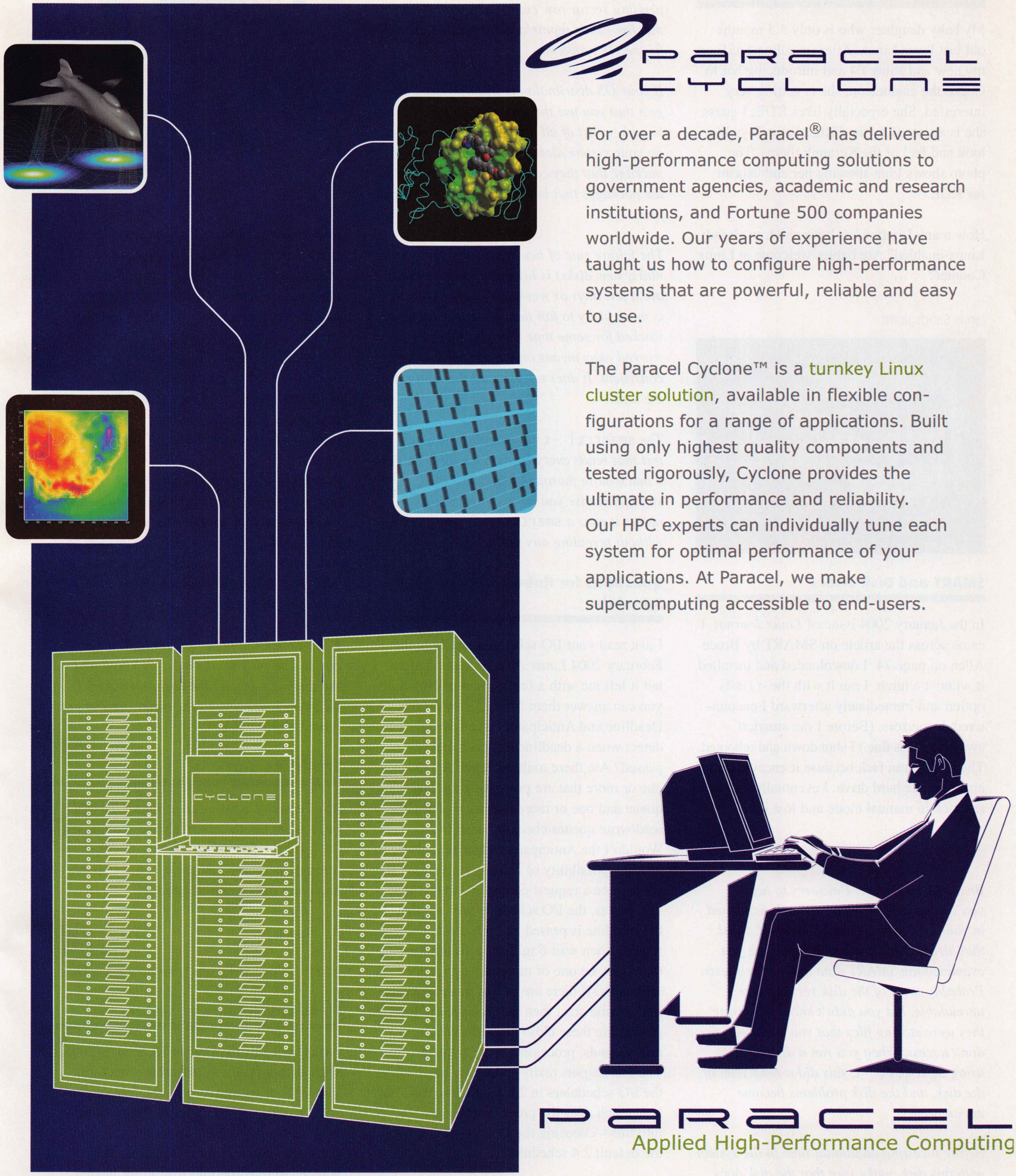
***LINUX JOURNAL* IS PUBLISHED BY, AND IS A REGISTERED TRADE NAME OF, SPECIALIZED SYSTEMS CONSULTANTS, INC.**
PO Box 55549, Seattle, WA 98155-0549 USA • linux@ssc.com

EDITORIAL ADVISORY BOARD
Daniel Frye, Director, IBM Linux Technology Center
Jon "maddog" Hall, President, Linux International
Lawrence Lessig, Professor of Law, Stanford University
Ransom Love, Director of Strategic Relationships, Family and Church History Department, Church of Jesus Christ of Latter-day Saints
Sam Ockman, CEO, Penguin Computing
Bruce Perens
Bdale Garbee, Linux CTO, HP
Danese Cooper, Manager, Open Source Program Office, Sun Microsystems

SUBSCRIPTIONS
E-MAIL: subs@ssc.com • URL: www.linuxjournal.com
PHONE: +1 206-297-7514 • FAX: +1 206-297-7515
TOLL-FREE: 1-888-66-LINUX • MAIL: PO Box 55549, Seattle, WA 98155-0549 USA • Please allow 4–6 weeks for processing address changes and orders • PRINTED IN USA

USPS *LINUX JOURNAL* (ISSN 1075-3583) is published monthly by SSC, Inc., 2208 NW Market Street, Suite 407, Seattle, WA 98107. Periodicals postage paid at Seattle, Washington and at additional mailing offices. Cover price is $5 US. Subscription rate is $25/year in the United States, $32 in Canada and Mexico, $62 elsewhere. POSTMASTER: Please send address changes to *Linux Journal*, PO Box 55549, Seattle, WA 98155-0549. Subscriptions start with the next issue. Back issues, if available, may be ordered from the Linux Journal Store: store.linuxjournal.com.

LINUX is a registered trademark of Linus Torvalds.



Discover how we can configure a system to optimize the performance of your computationally demanding applications

PARACEL CYCLONE

For over a decade, Paracel® has delivered high-performance computing solutions to government agencies, academic and research institutions, and Fortune 500 companies worldwide. Our years of experience have taught us how to configure high-performance systems that are powerful, reliable and easy to use.

The Paracel Cyclone™ is a turnkey Linux cluster solution, available in flexible configurations for a range of applications. Built using only highest quality components and tested rigorously, Cyclone provides the ultimate in performance and reliability. Our HPC experts can individually tune each system for optimal performance of your applications. At Paracel, we make supercomputing accessible to end-users.

PARACEL

Applied High-Performance Computing

Contact us today for your free expert consultation at 877.321.FAST www.paracelcyclone.com

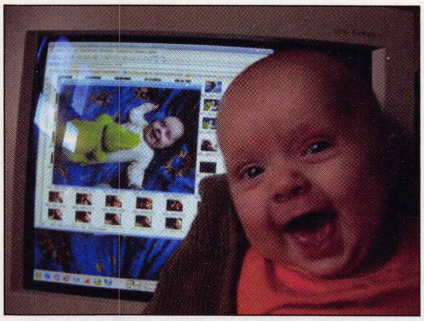
KDE Baby

My baby daughter who is only 3.5 months old just loves Linux. After installing SuSE on my new and shiny P4 and introducing her to Geeko the chameleon, she is getting very interested. She especially likes KDE. I guess she is attracted by the Fisher Price Toys-like look and feel of the Keramik theme. The photo shows Idun showing her enthusiasm for KDE.

How many Linux users have children already Linux-enabled? Are babies welcome at Linux Counter?

--

Janus Sandsgaard



SMART and Disk Errors

In the January 2004 issue of *Linux Journal*, I came across the article on SMART by Bruce Allen on page 74. I downloaded and installed it without a hitch. I ran it with the -t (test) option and immediately afterward I encountered disk errors. (Before I ran smartctl everything was fine.) I shut down and rebooted. The system ran fsck because it encountered errors on the hard drive. I eventually had to run fsck in manual mode and lost info.

--

Marshall Lake

***Bruce Allen replies:** I'm sorry to hear you've had disk problems. What's happened in your case is unfortunate. You had a disk that already was compromised before you examined the SMART data and ran self tests. Probably some of the disk sectors were unreadable, but you didn't know it because they were storing files that you normally don't access. When you ran a disk self test using* `smartctl -t`, *this did a read-scan of the disk, and the disk problems became apparent.*

Before investing additional time in the system with this disk, make sure that the disk does *pass a long self test without errors and that the Current pending sector and/or Offline pending sector raw counts are zero. If not, the disk needs repair/replacement before you fix the OS problems.*

If your OS distribution is RPM-based, I suggest that you use the verify options of RPM to make a list of all missing or changed files on your system along with the name of the package that they came from. Then, re-install the packages that have missing or corrupted files.

The failure rate of new hardware (and particularly new disks) is higher than when they are a few days or weeks old. New hardware is more likely to fail than hardware that has worked for some time. The fact that fsck worked okay means only that the filesystem is consistent. It does not *read all the data from the disk.*

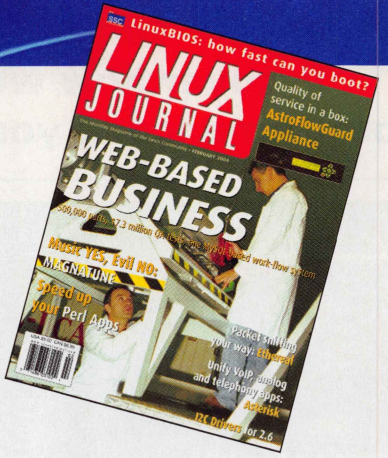
The `smartctl -t` *command starts a self test that reads* every *byte on the disk. This is a much more thorough test than fsck. I strongly advise you not to trust the disk unless it runs a* `smartctl -t long` *test without revealing any problems.*

Questions for Robert Love on I/O Scheduler

I just read your I/O scheduler article in the February 2004 *Linux Journal*. Great article, but it left me with a few questions. Hopefully you can answer them for me. How do the Deadline and Anticipatory I/O Schedulers detect when a deadline has been met or passed? Are there multiple kernel threads, one or more that are processing the main queue and one or more that are walking the read/write queues checking deadlines? Wouldn't the Anticipatory algorithm also give the possibility of causing starvation? If we assume a request comes in and the deadline passes, the I/O scheduler will see that the deadline is passed and process the request, then wait 6 milliseconds thinking there may be one or more requests in the same area. If there are one or more requests in the same area, then will it basically loop processing these close requests, waiting 6 milliseconds, processing more requests, etc.? The test outputs really show how much better the I/O schedulers in 2.6 are than in 2.4 , but how much does the process scheduler play into this—choosing the O(1) scheduler over the default 2.4 scheduler?

--

Torin Ford


LinuxBIOS: how fast can you boot?
LINUX JOURNAL
Quality of service in a box: AstroFlowGuard Appliance
WEB-BASED BUSINESS
Music YES, Evil NO: MAGNATUNE
Speed up Your Perl Apps
Packet sniffing your way: Ethereal
Unify VoIP, analog and telephony apps: Asterisk
I2C Drivers for 2.6

***Robert Love replies:** Glad you liked the article. There are no threads involved. The I/O scheduler code is handled in two places: in the interrupt handler for the disk's driver and in the process context code of the process submitting the request. The code can detect a deadline easily, because the deadline-sorted lists are inherently FIFO. The request at the head of the list has the oldest expiration, so the list need not be walked—just peek at the head of the FIFO list.*

In reply to the question about whether the anticipatory algorithm will also lead to starvation, eventually another request will expire, and it will go service it. The anticipation heuristic is just used to decide where to seek to, it otherwise does not change the algorithm of the disk.

I don't think the better process scheduler comes into play whatsoever. We know why the times are better: the writes-starving-reads phenomenon, so the better numbers are readily explainable.

English Lesson

The February 2004 issue of *Linux Journal* once again has the awful phrase "comprised of" (on page 56). Can I ask you to grep through incoming manuscripts to catch this annoying solecism in the future?

--

A. T. Young

It'll be in our stylebook from now on.—Ed.

Flaky Entertainment System

I was on a Virgin plane in December 2003 with such a faulty system that something like 90% of us mostly saw the reboot console! Luckily, it was a night flight, otherwise the kids would have gone crazy, and there was no mention of Linux on the screen, so most people would not have associated it with our beloved OS!

CAN A SERVER PROVIDE SCALABILITY, HIGH AVAILABILITY, EASY INSTALLATION, SERVICE AND MANAGEABILITY, WHILE LOWERING YOUR OWNERSHIP COSTS?

Aspen Systems' Glacier Xeon™ Blade Server based on the Intel® Xeon™ processor delivers equivalent performance and capabilities of a rack server, but with easier installation, service and management.

Aspen Systems' Glacier Xeon Blade Server System makes networking and storage expansion simple. The blade server platform is also highly reliable thanks to built-in redundancy for every component. No tools or expertise are required to add or replace a new blade in the **Aspen Systems**' Glacier Xeon Blade Server. The blade system eliminates excessive cabling and you can flexibly mix and match a wide variety of blades in the same chassis. Individual blades are easily managed locally or anywhere on your network with the Intel Server Management Module.

For additional information about **Aspen Systems**' Glacier Xeon™ Blade Server, contact **Aspen Systems**' team of sales engineers by calling toll-free (U.S. and Canada only) at **800.992.9242** or **303.431.4606** or via the Internet at **www.aspsys.com**

intel inside XEON

Aspen Systems inc.

THE BLADE SYSTEM ARCHITECTURE

- 7U Chassis supports up to 14 individual blades, each blade consisting of processors, memory and storage
- Dual Intel® Xeon™ processors running at 2.67 GHz, 2.8 GHZ, and future speeds
- Dual integrated network environment support
- Internal and external storage expansion capabilities
- Integrated Chassis-level system management and KVM switch for compute blade configuration, control and status

However, it also shows that we need to release only well-tested stuff to real customers, a lesson we need to learn if Linux is to move into the home environment.

--
Kirk Martinez, University of Southampton

Compatibility List?

Recently I went looking for a Linux compatability list on the Web and I was unable to locate one. Does *LJ* know where one is located?

--
Walt L. Williams

Every major Linux distribution maintains a hardware compatibility list. Not all distributions include the same hardware support, so you need to check yours. If you want to know whether there's a driver available in source code form that you can compile and configure, you usually can do best by searching for the hardware name +linux on Google.—Ed.

Table Fix for Work-Flow Article

I noticed a mistake in the February 2004 issue of *LJ* in the article by myself and Luciano Barone about REDACLE. In the text there is a reference to Table 2 that is missing. Table 2 should, in fact, read as follows:

ID	Name	Subname	Type
195	crystal	Barrel	1L
36	capsule	Barrel	C1
35	apd	Barrel	

Only the first row is relevant for text comprehension. Tables actually marked as 2 through 6 should be renamed 3 to 7.

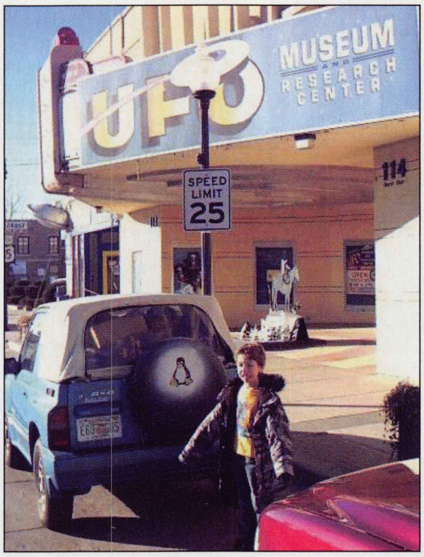
--
Giovanni Organtini

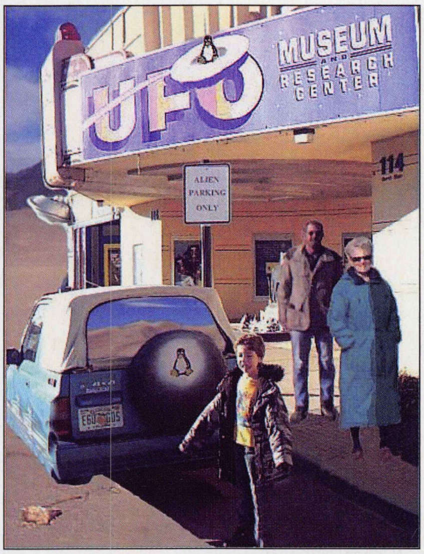
Photo of the Month

Here are a couple of photos of our last vacation. We went to Roswell, New Mexico to find out if the aliens were using Linux. Most were Debian supporters. There is a before and after photo. Some help from The GIMP was required.

--
Tom and Orion Maier

Photo of the Month gets you a one-year subscription. Send to ljeditor@ssc.com.—Ed.





Lindows: Good install, but Support Needs Work

After reading Steve Hastings' review of Lindows 4.0 [*LJ*, December 2003], I had a few comments of my own to add. First, concerning the Click-n-Run warehouse, I had many of the same complaints as Steve. Along with the rebranding of many packages and the inability to get packages to work properly, there is a serious problem with software versions being behind the curve. For instance, their OpenOffice.org version is 1.0.3, whereas the current OpenOffice.org is 1.1.x. Evolution in the warehouse is on version 1.2, while Ximian is on version 1.4. There are many examples of this.

Another serious problem is in the area of customer support. I sent a report to Lindows tech support in late December 2003 about these issues and have yet to receive anything other than an automated response ("We'll reply to this issue in a few days"). Since that initial report, I've sent two follow-ups, and those have gone unanswered as well.

Having said all that, I will agree that Lindows is very simple to install. I tried both 4.0 and the new 4.5 version, and they install in less than 15 minutes. There were problems with the video card (an ATI All-in-Wonder) and the media-card reader, but I have less problems with getting Lindows to work with the hardware than I did with Mandrake, Red Hat and that other company from Redmond.

But, until they have current version applications available in the Click-n-Run warehouse and until they resolve the issues with customer support, I won't be recommending Lindows to anyone.

--
Ed Dulaney

Big News at Linux Vendors?

The Linux magazines have been conspicuously quiet about two of the biggest Linux stories of 2003. The first is Red Hat rolling their Linux into Fedora Core and dropping support of their older versions. The second is Novell buying SuSE and Ximian. Red Hat is putting a positive spin on their story, while many others believe it's a blunder on Red Hat's part and a betrayal of their customers. Novell, is now a major Linux player. Does it all of a sudden have Linux religion? Will they contribute to Linux's evolution? Or is Linux just the latest strategy/weapon in their age-old war with Microsoft? Both these stories have a big impact on the Linux community, and they are getting a lot of coverage and discussion on-line. I would expect *LJ* to join in on the same. Will this happen soon?

--
Henry E. Alubowicz

See the next letter for why these stories might be less important than you think.—Ed.

LJ Discovers IT's "Dark Matter"

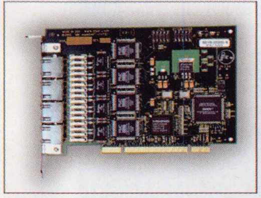
Just wanted to commend Doc on a great piece ("DIY-IT..." [*LJ*, February 2004]) and *LJ* for being a brave enough to print it, despite its vendor-specific (read: advertiser) implications.

LETTERS CONTINUED ON PAGE 81

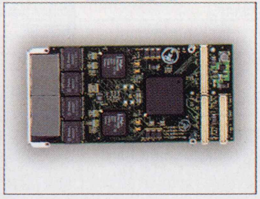
Which direction?

You decide...Linux ready on every board

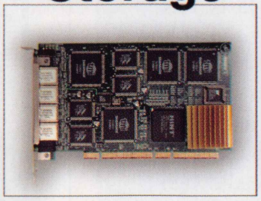
WAN



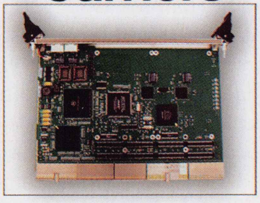
LAN



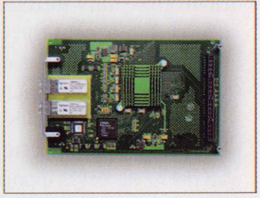
Storage



Carriers



Custom



SBE®

Linux On Demand

flexibility on demand | 925-355-2000 | info@sbei.com | www.sbei.com

diff -u: What's New in Kernel Development

The **2.6.0** kernel has come out, roughly on schedule. The arrival of a stable series at a predicted time is a first in Linux development, and it's none too common in the rest of open-source development either. Although most people tend to accept **Linus Torvalds'** discovery that encouraging help from arbitrary developers actually does result in better code and faster development, the secret of predictable release cycles has remained elusive.

In the case of the 2.6.0 release, Linus did insist on maintaining feature freezes and other policies without going back on them, but this itself could work only if the various features were made to coalesce at roughly the same time. Clearly, no simple method has emerged yet. In terms of maintainership of the 2.6 tree, this job apparently has gone to **Andrew Morton**, in much the same way that the 2.4 tree went to **Marcelo Tosatti** and the 2.2 tree went to **Alan Cox**. Actually, even now it is unclear who is the true maintainer of the 2.6 tree. Andrew certainly appears to be making decisions with authority, but Linus is the one who puts out actual releases. Perhaps Linus and Andrew will continue in tight collaboration until the 2.7 fork, but for now the overall maintainership situation is much less clear than it was for any of the previous stable series. It will be interesting to see how these issues play out in the 2.7–2.8 time frame.

Last month I said **XFS** might not make it into the 2.4 tree due to Marcelo Tosatti's plans for a deep freeze following the emergence of the 2.6 series. In fact, XFS did make it into 2.4 before the final cutoff. This was partly due to a large developer outcry, in which it was pointed out that XFS was the last major journaled filesystem missing from 2.4. In addition, Marcelo himself previously had indicated that XFS would go into his tree when it was ready. In the end, Marcelo did insist on a thorough analysis of the code by a third party before allowing it into 2.4. Ultimately, however, XFS made the cut, and further XFS work is going into 2.4 to round off the feature in spite of the deep freeze. Undoubtedly, this will taper off quickly.

Ian Kent has (for the moment at least) accepted the role of **DevFS** maintainer. This is a sharp reversal of the previous trend, which was to let DevFS die a quiet death and replace it with a similar feature, such as **udev**. In fact, **Greg Kroah-Hartman** has been making regular udev releases, with an eye toward exactly this possibility. However, in spite of DevFS' flaws, including some that are apparently so severe as to be virtually unfixable, DevFS still provides features that no other system has yet been able to surpass. Folks like Ian have begun to advocate keeping DevFS in the kernel and fixing whatever problems it may have, however difficult they may turn out to be. Ian, for as long as he lasts, has decided to accept the challenge. It will be interesting to see how the other kernel developers, notably **Alexander Viro** (a particularly vehement DevFS-hater), take to this unexpected turn of events.

Yasunori Goto has been working on some code to allow **RAM hot plugging**. So far he's been targeting NUMA machines, but high on his to-do list is a port to **IA-64** and **IA-32** machines. Although interesting for desktop users, this feature could be quite a boon to server administrators, who often are blamed for downtime on company systems. It also brings us closer to truly stackable systems. Imagine a wearable computer that gains two additional processors when you put on your shoes. Yasunori may not be thinking quite along those lines, but the prospect of hot-plugging RAM chips is enticing.

—ZACK BROWN

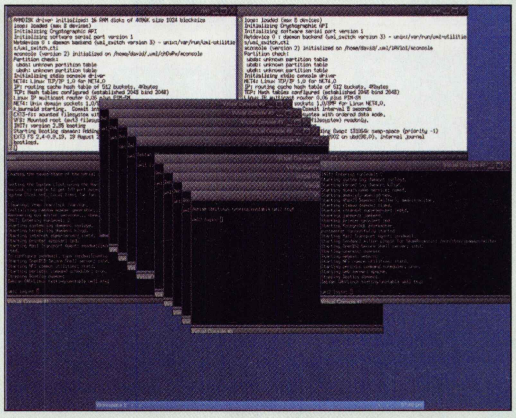
The Godfather of SOLIS

Jon "maddog" Hall and Cesar Brod first met in 1999. In 2001, he visited Univates, where he actually suggested the team should think of spreading what they were doing to other universities. He is considered to be the Godfather of SOLIS. See page 96 of this issue for more information about SOLIS, the Brazilian Free Software Cooperative.

—CESAR BROD

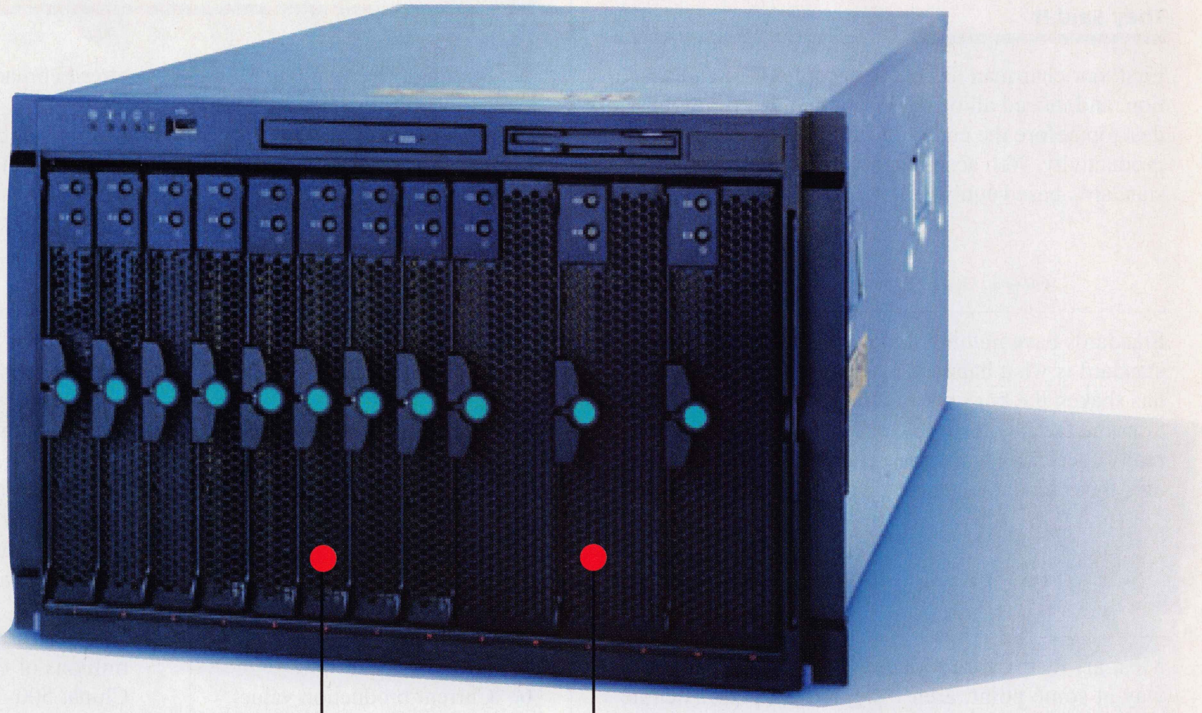
User-Mode Linux: user-mode-linux.sourceforge.net

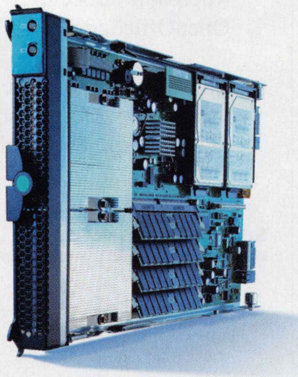
Definitely not for the faint of heart or for RAM/CPU-challenged systems. User-Mode Linux (UML) allows you to build running Linux systems within Linux systems for use as sandboxes or virtual servers. It's a poor man's way to emulate a mainframe running Linux instances. You can bang on your sandbox all day without affecting your host. Perfect for setting up lab rats, honeypots or verifying that the latest updates won't hose your other systems. If you made a copy of the file beforehand, going back is all too simple. It's also great for hosting providers to give clients their own servers without the additional hardware costs. Requires: running Linux system with X (preferably with the optional skas3 patch) and a uml-patched Linux kernel.



—DAVID A. BANDEL

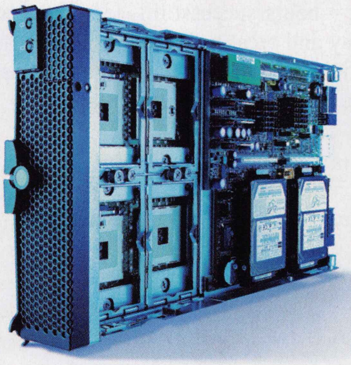
Density redefined.





2-Way Intel® Xeon™

Up to 14 Compute Blades



4-way Intel® Xeon™

Powered by the Intel® Xeon Processor MP

Check your physics books, density now = Q7. The Q7 Server Blade Chassis is innovative blade engineering at its best. It has the capability to increase processing power while decreasing space and management resources. Accommodating 2-Way and 4-Way Server Compute Blades containing Intel® Xeon® processors MP, all nicely enclosed in 7U. While QSOL is rewriting the definition of density, log on to Qsol.com and order one today!

QSOL.COM is a trademark of QSOL.COM INC. Intel is a registered trademark of Intel Corporation.

QSOL.COM
Server Appliances℠
800.933.7510

They Said It

First, our chairman has challenged the IT organization, and indeed all of IBM, to move to a Linux-based desktop before the end of 2005. This means replacing productivity, Web access and viewing tools with open standards-based equivalents.

—BOB GREENBERG, CIO, IBM
(www.theinquirer.net/?article=13485)

Standards have nothing to do with innovation; a good standard is what happens when an industry basically has shaken the bugs out of a technology and then, after the fact, writes it down. This is true of all the really successful standards: grams and meters, voltage, the calendar, octane ratings, TCP/IP, XML.

—TIM BRAY
(www.tbray.org/ongoing/When/200x/2003/05/10/RSS-std)

Most art forms are produced "live" in the same way at some point, aren't they? If art patrons came to pay for the performance, and being involved in the performance, rather than for the artifact itself as if it was a bar of gold, then P2P wouldn't be the threat it's characterized as.

This is already what's happening with software and open source—being successful is less about hiding yourself away and creating the perfect code, and more about participating in the community process by which a body of communal code is written.

—BRIAN BEHLENDORF (FROM AN E-MAIL)

The world rewards action. It doesn't reward much of anything else.

—SCOTT ADAMS,
(paulboutin.weblogger.com/2003/12/22)

convertfs: tzukanov.narod.ru/convertfs/index.html

Need to convert from ext2 to something like ReiserFS or XFS? With convertfs, you can do it on the fly on any unmounted filesystem or on one you can unmount from a running system. Don't try this at home—or anywhere else for that matter—without a backup. However, I successfully converted several ext3 filesystems to XFS. The included bash script even reminds you to change your /etc/fstab entry for that partition. Requires: glibc and kernel with loop device support.

—DAVID A. BANDEL

LJ Index—April 2004

1. Starting price in dollars for a Linux PC by CPUBuilders at Sam's Club: 256.68
2. Starting price in dollars for a Linux PC by Microtel at Wal-Mart: 199.98
3. Price in dollars for a loaded 2.66GHz Linux-Certified Debian Linux Laptop: 1,749
4. Billions in dollars Japan is spending to fund Linux developers in Japan: 8.3
5. Number of local companies in Taiwan involved in open-source software development: 20
6. Current production value, in millions of dollars, of open-source software developed in Taiwan: 3.4
7. Millions of dollars Taiwan government will put into promotion of open-source software development: 3.4
8. Percentage of servers in Taiwan currently running open-source software: 10
9. Goal percentage for servers in Taiwan running open-source software by 2007: 30
10. Percentage of personal computers in Taiwan currently running open-source software: 0.2
11. Goal percentage for personal computers in Taiwan running open-source software: 5
12. Stock value in billions of dollars paid by Sun Microsystems for Cobalt Networks in 2000: 2
13. Number of Cobalt products Sun continues to make after February 19, 2004: 0
14. Millions of copies of the Linux-based Sun Java Desktop System projected for installation in China: 200
15. Minimum starting install rate per year for Sun Java Desktop Systems, beginning January 2004, in thousands: 500
16. Maximum starting install rate per year in millions for Sun Java Desktop Systems, beginning January 2004: 1
17. Sun's goal number in millions of Linux PCs in China: 500
18. Range of percentage of cost reductions made possible by migrating from Microsoft Office to OpenOffice.org: 60–70
19. Thousands of citizens per week using 72 open-source telecenters in São Paulo: 150

Sources

1: Sam's Club

2: www.walmart.com

3: Linux Certified

4: itbusiness.ca

5–11: *Asia Computer Weekly*

12, 13: *CXO Today*

14–16: *Motley Fool*

17: John Fowler of Sun, at Apachecon

18: John Terpstra

19: Nat Friedman

FANATICAL SUPPORT™ IS THE DIFFERENCE

Rackspace — Managed Hosting backed by Fanatical Support.™

Servers, data centers and bandwidth are not the key to hosting enterprise class Web sites and Web applications. At Rackspace, we believe hosting is a service, not just technology.

Fanatical Support is our philosophy, our credo. It reflects our desire to bring responsiveness and value to everything we do for our customers. You will experience Fanatical Support from the moment we answer the phone and you begin to interact with our employees.

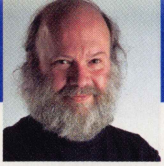
Fanatical Support has made Rackspace the fastest-growing hosting company in the world. Call today to experience the difference with Fanatical Support at Rackspace.

LINUX JOURNAL READERS' CHOICE 2003

Thanks for
honoring us with the
2003 Linux Journal
Readers' Choice Award for
"Favorite Web-Hosting Service"

rackspace™
MANAGED HOSTING

1.888.571.8976 or visit us at **www.rackspace.com**



Ten Years of *Linux Journal*

They said he was crazy when he asked about device driver support. But look at Linux now.

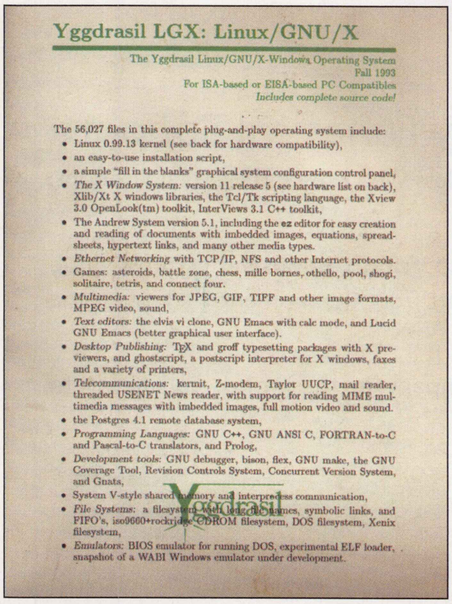
BY PHIL HUGHES

Is she still a Linux gamer?

With this issue, *Linux Journal* turns ten. I hadn't really thought about how *LJ* had been my job for ten years until Don suggested I write this editorial. To me, that's a good thing; I have been publishing rather than counting days. What has happened in these last ten years is amazing. Ten years ago I don't think anyone would have expected to see the L-word in ads by HP, IBM, Oracle and many other big players.

In looking around to see what I had from ten years ago that might have something to do with this editorial, I found two interesting things. The first was my copy of *Yggdrasil LGX: Linux/GNU/X*. This is the Fall 1993 distribution. The second was a picture of Phaedra, the daughter of Joanne Wagner, our first ad rep. The two go together because Phaedra (now 15) used to play the text game *Mille Bornes*, from the BSD games collection, on this version of Linux.

The distribution came in the form of a 64-page book and had a description of what it included and what it would run on on the covers. Inside you found a CD, a regular 5¼" boot floppy and one of those new-fangled 3.5" boot floppies. The system requirements were 4MB of RAM and from 2 to 680MB of disk space. Inside the book were installation information and a list of where to get support—12 places. One of these 12 is Russ Nelson at Crynwr Software, who still is very much an active member of what we were all calling the free software community back then.



Yggdrasil LGX: Linux/GNU/X

The Yggdrasil Linux/GNU/X-Windows Operating System
Fall 1993
For ISA-based or EISA-based PC Compatibles
Includes complete source code!

The 56,027 files in this complete plug-and-play operating system include:

- Linux 0.99.13 kernel (see back for hardware compatibility),
- an easy-to-use installation script,
- a simple "fill in the blanks" graphical system configuration control panel,
- *The X Window System:* version 11 release 5 (see hardware list on back), Xlib/Xt X windows libraries, the Tcl/Tk scripting language, the Xview 3.0 OpenLook(tm) toolkit, InterViews 3.1 C++ toolkit,
- The Andrew System version 5.1, including the **ez** editor for easy creation and reading of documents with imbedded images, equations, spreadsheets, hypertext links, and many other media types.
- *Ethernet Networking* with TCP/IP, NFS and other Internet protocols.
- Games: asteroids, battle zone, chess, mille bornes, othello, pool, shogi, solitaire, tetris, and connect four.
- *Multimedia:* viewers for JPEG, GIF, TIFF and other image formats, MPEG video, sound,
- *Text editors:* the elvis vi clone, GNU Emacs with calc mode, and Lucid GNU Emacs (better graphical user interface).
- *Desktop Publishing:* TeX and groff typesetting packages with X previewers, and ghostscript, a postscript interpreter for X windows, faxes and a variety of printers,
- *Telecommunications:* kermit, Z-modem, Taylor UUCP, mail reader, threaded USENET News reader, with support for reading MIME multimedia messages with imbedded images, full motion video and sound.
- the Postgres 4.1 remote database system,
- *Programming Languages:* GNU C++, GNU ANSI C, FORTRAN-to-C and Pascal-to-C translators, and Prolog,
- *Development tools:* GNU debugger, bison, flex, GNU make, the GNU Coverage Tool, Revision Controls System, Concurrent Version System, and Gnats,
- System V-style shared memory and interprocess communication,
- *File Systems:* a filesystem with long file names, symbolic links, and FIFO's, iso9660+rockridge CDROM filesystem, DOS filesystem, Xenix filesystem,
- *Emulators:* BIOS emulator for running DOS, experimental ELF loader, snapshot of a WABI Windows emulator under development.

Who came up with the name?

We publish a magazine, so let me look at what I predicted back in the beginning. My editorial in issue two was a piece of fiction describing what Linux would be like in the year 2000—six years into the future. The first sentence says, "In the past 7 years we have seen Linux go from an idea for a small UNIX-like system into a movement to bring affordable, reliable multi-tasking software to anyone who could buy a rather minimal computer." I don't think there is any argument there.

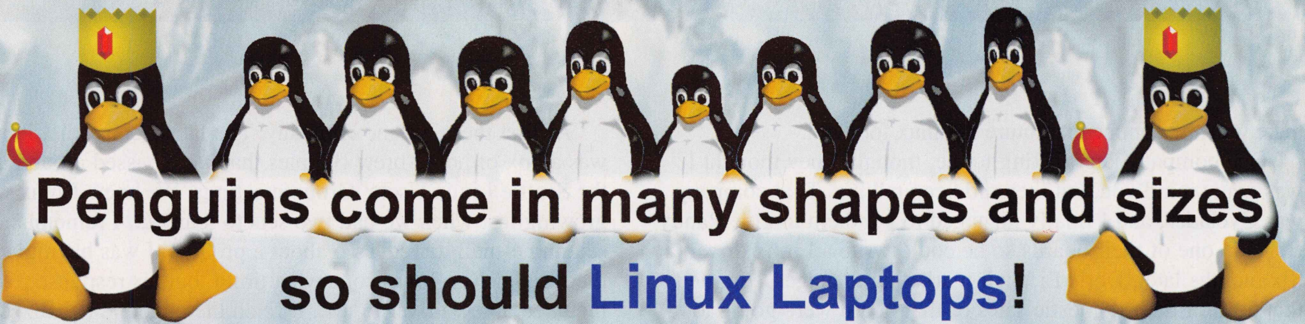
After some rambling about a program loader called MS-DOS, I went on to say, "With the advent of ISDN in the early 1990s and personal satellite stations in the late 1990s, connectivity became the big issue." Personal satellite certainly did happen. I was wrong about ISDN (I guess I forgot it stands for It Still Does Nothing), but DSL and cable clearly filled that gap. So, I am still on track.

And that's where the track went astray. For example, I predicted that 90% of *LJ* subscribers would be receiving the magazine on-line. It still sounds like a great goal, but a combination of people wanting to have something to carry on the bus with them and the way subscription audits work—that is, only paper magazines count toward the official circulation—has slowed progress there. Of course, to my credit, Microsoft founder Bill Gates previously had predicted that Xenix on an Intel 80286 chip was the future of computing, so at least I was a little closer.

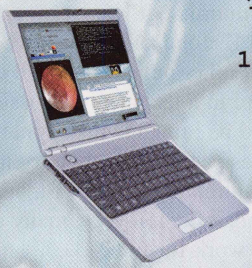
All my other predictions had to do with getting everyone on the Internet. In January 2001, I moved to Costa Rica. In the 1994 editorial I claimed that in 2000, I was in Yaak, Montana. So, I did move but I picked a place with a lot better weather. This move also helped me adjust my perspective about Internet connectivity. While many countries, Korea being a good example, are delivering broadband Internet service to a large percentage of their population, many other places are still without.

About nine years ago we decided we needed a Web presence. Linux and Apache sounded like the right approach, so we set up a 486DX100 with 16MB of RAM to test the waters. We agreed to evaluate what we really wanted to do when we got to 10,000 hits per month. *LJ* was growing and we were busy with other projects. When we finally looked at the Web site again we were at 100,000 hits per month. The system was handling it just fine. Today, we receive over 10,000 hits per hour on our Web sites.

Early on in the life of Linux, ISPs considered it to be an alternative to proprietary UNIX platforms. This was an era where dial-up was almost always the answer. Unfortunately, no intelligent serial communications boards were available with Linux drivers. I started talking to vendors and they all thought I was crazy to think

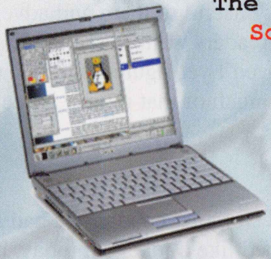


Penguins come in many shapes and sizes so should Linux Laptops!



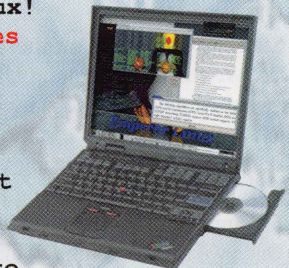
The **Meteor**: 2 lb Linux!
Sharp Actius MM10
1.0 GHz Transmeta Crusoe
15 GB HD
256 MB RAM (Fixed)
802.11b Wireless
10/100 Ethernet
ACPI Hibernate
USB Cradle
0.5" Thin!

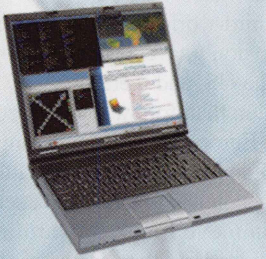
The **Raven**: 3.5 lb Linux!
IBM ThinkPad X Series
1.3-1.6 GHz P-M
20-40 GB HD
256-2048 MB RAM
10/100/1000 Ethernet
802.11b/g Wireless
FireWire
1" Thin!
APM Suspend/Hibernate



The **BlackPerl**: 4 lb Linux!
Sony Vaio V505 Series
2.0-2.4 GHz P4-M/
1.4-1.7 GHz P-M
40-80 GB HD
256-2048 MB RAM
CDRW/DVD
FireWire
802.11b/g Wireless
ACPI Hibernate

The **Toucan**: 5 lb Linux!
IBM ThinkPad T Series
X@1400x1050
1.3-1.7 GHz P-M
30-80 GB HD
256-2048 MB RAM
10/100/1000 Ethernet
802.11b/g Wireless
CDRW/DVD
APM Suspend/Hibernate





The **Gazelle**: 8 lb Linux!
Sony VAIO GRT Series
X@1600x1200
2.2-2.8 GHz P4-M/
3.06 GHz P-4
40-80 GB HD
256-2048 MB RAM
ACPI Hibernate
802.11b/g Wireless
CDRW/DVD or DVD+/-RW

The **Rhino**: 7 lb Linux!
Dell Latitude D800
X@1920x1200
1.6-1.7 GHz P-M
30-80 GB HD
256-2048 MB RAM
CDRW/DVD, DVD, CDRW,
or DVD+/-RW
802.11b Wireless
10/100/1000 Ethernet

Since 1999, EmperorLinux has provided pre-installed Linux laptop solutions to universities, corporations, and individual Linux enthusiasts. We specialize in the installation and configuration of the Linux operating system on a wide range of the finest laptop and notebook computers made by IBM, Sony, and Dell. We offer a range of the latest Linux distributions, as well as Windows dual boot options. We customize each Linux distribution to the particular machine it will run upon and provide support for: ethernet, modem, wireless, PCMCIA, USB, FireWire, X-Server, CD/DVD/CDRW, sound, APM, and more. All our systems come with one year of Linux technical support by both phone and e-mail, and full manufacturers' warranties apply. Visit www.EmperorLinux.com or call 1-888-651-6686 for details.

Successful Linux laptop solutions found for:

Sandia Nat Lab DOE NASA JHU Applied Physics Lab Max Planck Inst Woods Hole NRAO Naval Research Lab Smithsonian Institute Los Alamos Nat Lab UC Lawrence Berkeley Labs NOAA Lawrence Livermore Nat Lab US Army Research Lab U Penn U Utah U Maryland Yale Cornell Columbia U Oregon NC State Duke U Massachusetts George Mason Colorado State Caltech U Illinois Georgia Tech U Minnesota U Chicago Pace U UCSF UCSD Rutgers Princeton NYU Purdue Boston U MIT U Arizona George Washington U U Hawaii U Maine

www.EmperorLinux.com **1-888-651-6686**

Model prices, specifications, and availability may vary. All trademarks are the property of their respective holders.

there was any commercial future in Linux for them.

One company kept talking to me, though. They thought I was crazy mind you, but they did keep talking. That company was Cyclades, and I finally managed to get them to give Randy Bentson one of their boards so he could write a Linux driver. Six months later, Doris Li, their marketing manager, admitted that 50% of their domestic sales of that board were going to Linux users. Much like Russ Nelson, Cyclades is still here.

What We Did Right (and Wrong)

When Irene Pasternack and I started SSC about 20 years ago, we knew we would be doing computer-related documentation and training, but we needed to focus better than that. After discussing this, we decided the important thing that would make us different from everyone else was we would create only products we wanted for ourselves.

We weren't perfect, and every now and then we would get a brilliant idea and forget to check it against this criteria. For example, we did three MS-DOS Reference Cards. Bottom line: we sold about 5,000 of them, while we sold more like 500,000 vi Reference Cards.

Starting *LJ* was one of those things we wanted for ourselves. We had been running UNIX, but it looked like Linux was what we needed. We started the magazine and made the switch pretty much at the same time and have never looked back. We still do layout using Quark XPress on a non-Linux system, but Scribus is looking pretty good now, so we could be an all-Linux shop in the near future.

Once we started doing *LJ*, things have fallen into place a lot better than ever before. Here are a few of those magic happenings:

- I knew Doc Searls' wife before they were married, in fact, before he knew her. Doc became part of *LJ* because Joyce convinced him that I seemed to be on to something.

- Right after Doc wrote the *Cluetrain Manifesto*, he got e-mail from Dawn Smith. She told him she really liked the book and if he was ever in Costa Rica....He wrote back telling her that his publisher at *LJ* had just been to Costa Rica. Dawn told him that her husband recently had been working with Linux. Today, Willy Smith, Dawn's husband, works for us.

- Back when I was starting SSC I worked for a company that did point-of-sale systems for gas stations. That's where I met Dan Wilder, who has been our head geek for the last few years.

I could go on and on here, but the important point I am trying to make is that with *LJ*, connections have just happened. Or, more accurately, when we let things happen in the Linux racket we seem to get the desired results.

Building a Community

Linux made me believe in community. I have worked in computers since 1968 and had a few other technical jobs when I was going to college. They were just that, jobs. I worked for a company that produced something it wanted to sell. You always had "the company" and "the customer". With Linux, that paradigm changed.

At first there was no company. The latest Linux distribution was a box of home-brew floppies that were passed around at the Seattle Linux Users' Group meetings. We knew Linux was growing fast when we had two sets to loan out. I remember sending e-mail to Ted Ts'o about a problem I was having with the serial driver. I was somewhat timid, but his response included a new driver to test. It fixed the bug, and I realized that Linux was moving forward because we worked together.

The first time I met Linus Torvalds I saw what was behind that whole sense of community. It was at a party in Washington, DC. When Michael K. Johnson and I arrived, Linus was there along with a few other people who had contributed code to Linux. They were in a technical discussion about how something new should be implemented. What I saw was Linus treating these people as peers rather than trying to be the boss. To me, Linus is the ultimate manager. I mean, who else has been able to get thousands of employees, many of whom he has never even met, to work for free?

With Linux and *Linux Journal*, that community continued. We have authors who write because they want to write, readers who tell us that our ads are very useful to them and advertisers who ask us what we think our readers want to see. Sure, there is money changing hands (I haven't yet figured out how to get my staff to work for free), but much of that money is recycled within the Linux community.

World Domination

Linus coined the idea of World Domination after we had started *Linux Journal*. At first it sounded like a joke, but today, it sounds like a goal we will reach before *LJ* turns 20. But, again, looking at the world scene we see a much different picture than if we solely focus on the United States.

Initially, I thought it was a third-world issue that drove Linux penetration. That is, people here don't have the same amount of disposable income as those in the US, and therefore, they are more willing to listen to a solution than follow the marketing hype. But, Linux penetration in Europe is very significant; the same is true in Asia. Although I don't understand why, I think the United States will be the last country to take Linux seriously.

There are two ways we can move Linux forward in the marketplace. One is to continue to show the shortcomings of proprietary alternatives. The problem with this approach is you get into a "which is better" contest, and the people on the anti-Linux side have a lot more money than we do. Thus, it isn't an issue of being right but of getting everyone else to see you are right.

The other method is to just do it. That is, run Linux. Help others run Linux. When a business needs a solution, offer Linux. And, most important of all, when Linux doesn't do something that is needed, address it. We have an amazing base available and some amazing talent in the Linux community. It is time to run with what we have and reach that World Domination goal.

What's Next?

First, I have some questions that need answers. Actually, we probably will never get the answers, but I find them interesting to think about:

- At what point did Microsoft spend more on bad-mouthing Linux than all Linux vendors combined spend on marketing?

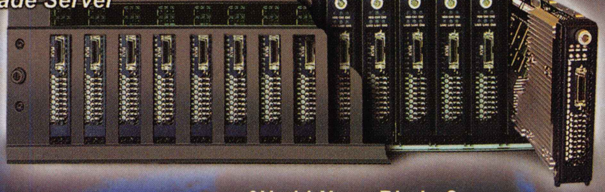
High-Performance Blade Servers

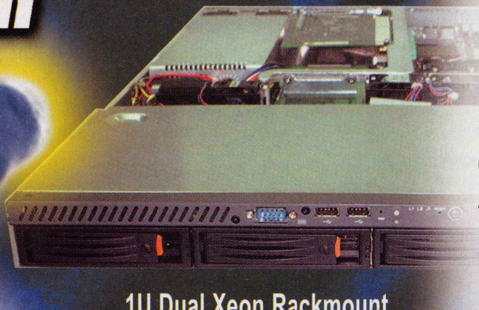
2U- 16 PIII Blade Server

Low-Power Consumption

PRODUCT EXCELLENCE AWARDS
LINUXWORLD CONFERENCE & EXPO

2003 Best System Integration Solution
3U Blade Server


3U- 14 Xeon Blade Server


1U Dual Xeon Rackmount

Point to
tsti.com

Choose from a broad line of Tatung rackmount servers.

Blade Servers
Superior Tatung architecture integrates 16 or 14 server blades in a 2U or 3U chassis. That form factor can deliver 336 PIII or 196 Xeon servers in a single 42U height rack while using significantly less power than comparable systems and eliminating more than 80% of the wires.

Each hot-swappable blade is powered by the latest low-power Intel® ULV PIII or LV Xeon processor and has built-in dual gigabit ethernet switches and advanced management software.

1U Xeon Servers
High Performance single or dual Intel Xeon processor rackmount servers are now shipping. Configure the 1U with up to four hard drives and a 2U or 3U with six hard drives or more.

Tatung is building rackmount servers now. Call today and have us build a system for you.

The Intelligent Choice in Server Solutions

TATUNG®
SCIENCE & TECHNOLOGY, INC.

USA: 1-800-659-5902 • www.tsti.com
Europe: www.tatung.nl
Asia: www.tatungserver.com

Intel is a registered trademark of Intel Corp.
Tatung is a registered trademark of Tatung Company.

- When will the number of installed copies of Linux exceed the number of legally installed copies of Microsoft OSes?
- When will the number of installed copies of Linux exceed the number of total installed copies of Microsoft OSes?

Notice that I say when, not if. This will happen. It may happen last in the United States, but the combination of Linux maturing and the world economy dictates that this will happen. All of us have an opportunity to make this happen faster.

By the time you read this, I probably will be in Nicaragua helping get some Linux classes started. These are not classes for systems administrators or even computer users. The average person will have never used a computer before and currently makes about $3/day working in a cigar factory. This is not the profile of people that Microsoft is interested in marketing to, but they are typical future new computer users. This is my current path; you don't have to pick this same path. You simply can use Linux and set an example. Another good approach is to help someone convert to Linux. For example, my neighbor does computer consulting. All his customers run Microsoft software, but they are getting pretty irritated by worms and viruses. Most of them run typical office software, including word processor, spreadsheet and e-mail, so it is the perfect time to offer them an alternative.

Although Linux certainly has become mainstream, it still has a great distance to go before we can claim a World Domination success story. If you are a *Linux Journal* reader, you are likely ahead of the crowd. Put a little effort into getting that crowd moving in the right direction, and we can reach World Domination long before *Linux Journal* is 20.

Looking beyond Linux, we also need to look at how what has happened with Linux is a template for what can happen in other areas. In mid-2003, we started our WorldWatch site as an experiment (worldwatch.linuxgazette.com). We saw the need for a worldview of what was happening with Linux and open-source software from a social, political and economic point of view. What happened made WorldWatch Editor Willy Smith realize that we needed to tie together free, libre and open-source software (FLOSS) with similar efforts in other areas. One of the best examples is what we called Open Source seeds. That is, seeds that actually can reproduce rather than the genetically engineered ones that have to be purchased from the patent holder.

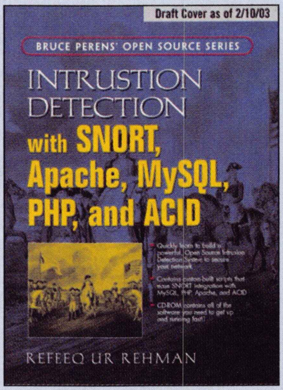
This revelation resulted in our creation of a new Web site, A42 (www.a42.com). The Linux-related technical side of what WorldWatch was will be appearing on Linux Gazette (www.linuxgazette.com). A42 is going to be where we try to tie together the whole open-source revolution—whether it applies to computers or not. In order to remain sane, A42 will take a light-hearted approach.

So, let's just say Linux moving toward World Domination is well on the way, and I feel comfortable enough it will happen that I am going to go further out on a limb this time and say that Open *Everything* is the real goal. *Hasta pronto.*

Phil Hughes is publisher of *Linux Journal.*

BOOK REVIEW

Intrusion Detection with SNORT, Apache, MySQL, PHP, and ACID


Draft Cover as of 2/10/03
BRUCE PERENS' OPEN SOURCE SERIES
INTRUSTION DETECTION
with SNORT, Apache, MySQL, PHP, and ACID
REFEEQ UR REHMAN

by Refeeq Ur Rehman

Prentice Hall PTR, 2003
ISBN: 0-13-140733-3
$39.99 US

Network security is a hot topic these days, and intrusion detection systems are playing a greater role in network security. Refeeq Ur Rehman's book *Intrusion Detection with SNORT, Apache, MySQL, PHP, and ACID* claims to explain and simplify all aspects of SNORT, from building to managing an intrusion detection system (IDS) in your network.

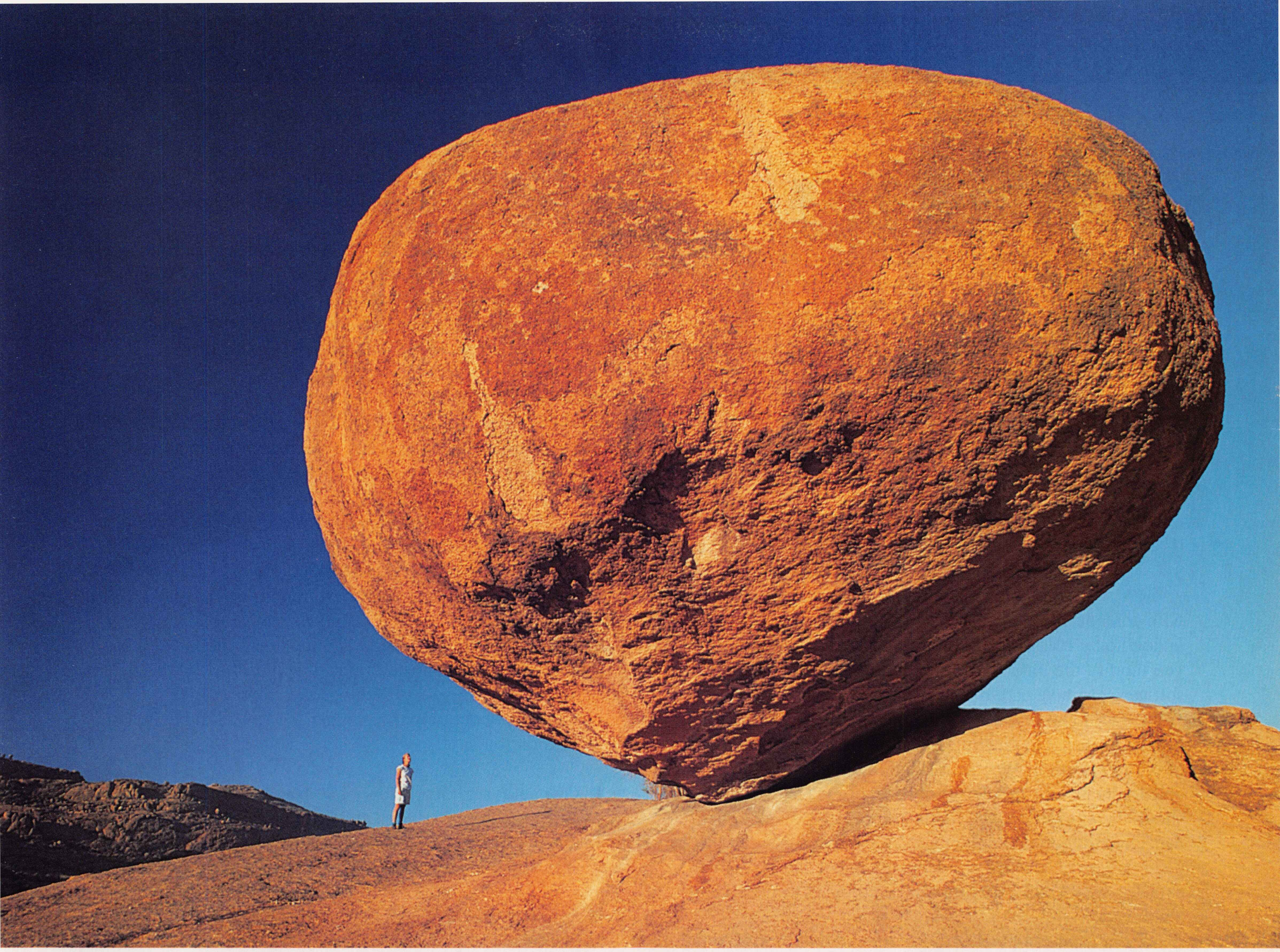
This book presents a front-to-back solution for building an IDS and includes quite a few example scripts and code snippets. It is a good technical treatment of installing and configuring SNORT and running the IDS in conjunction with Apache, MySQL, PHP and ACID.

The coverage of SNORT rules is helpful, because Rehman provides examples and an explanation of what makes a good or bad rule. He discusses several plugins and output modules and explains what each one does and how to integrate it into the IDS. Most interesting to me was the inclusion of a MySQL database for logging and the use of Analysis Console for Intrusion Detection (ACID) to create a Web interface. SnortSnarf is discussed as an alternative for presenting SNORT data over the Web. Barnyard, SnortSam and IDS Policy Manager are other useful tools covered in the text.

Most of the information pertaining to installing and using SNORT is readily available in the SNORT documentation or on the Web. Nonetheless, this book is a useful and handy reference.

I noted a few grammatical and technical errors in the book; however, they are unlikely to diminish its readability or value. Although the book covers SNORT version 1.9, much of Rehman's material should apply to version 2.0 as well. Appendices include a MySQL primer, an introduction to tcpdump and packet header formats.

—SEAN TIERNEY



Network backup seems insurmountable if you don't have the right solution.

Introducing Arkeia 5.2
The right solution.

If you're responsible for protecting your company's heterogeneous network, the Arkeia 5.2 Linux or Unix backup server solution is just what you've been looking for.

Arkeia has always been noted for powerful, scalable, reliable and cost-effective backup solutions. Now we've added even more options to give you greater control and ease of use.

New features, greater benefits

Our completely redesigned reporting functions allow you to produce highly customizable machine-based reports as well as more developed messages that inform you about successful operation completion.

Quantum DLTSage™ support for SDLT tape devices helps you predict and prevent storage system errors.

An optimized catalog algorithm lets you dramatically increase backup speed for large file servers—five times faster for 10,000 files within a single directory and up to 50 times faster for 150,000 files!

We've also expanded GUI capabilities for maximized security. You can easily clone tapes for off-site storage. Combine this with Arkeia's Disaster Recovery module, and you can automatically rebuild servers from scratch if they are destroyed or stolen.

In addition, a new hot backup plug-in is available for MS-Exchange that complements our online backup solutions for Oracle, Lotus, DB2 and MySQL.

Try Arkeia 5.2 for 30 days. Free!

The best way to prove that Arkeia 5.2 is the right solution for you is to let you try it—**FREE**—for 30 days. We'll even include free installation tech support. Simply download the demo version at www.arkeia.com.

arkeia®
ENTERPRISE NETWORK BACKUP

www.arkeia.com

COREBlog

It's not a Web development framework without a blog package. Here's a Zope-based system for a Web journal with all the bells and whistles.

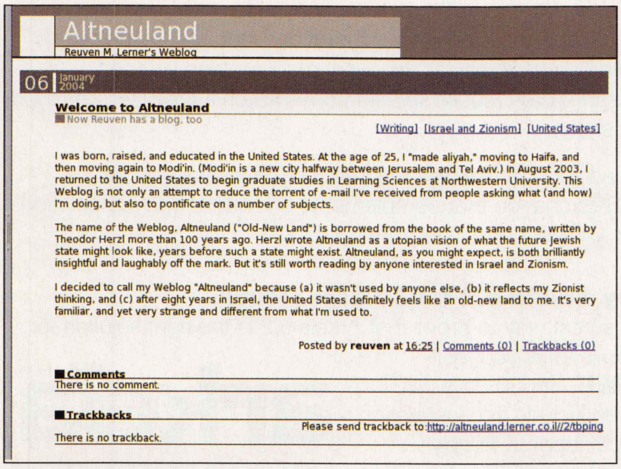
BY REUVEN M. LERNER

Over the past year or so, this column has looked at a number of open-source products that can serve as a content management system (CMS) for a Web site. If you publish a newspaper, magazine or Web site with content that changes on a regular basis, such a CMS undoubtedly could be a boon to your site. After all, why should you modify the links, headlines and other items by hand, if software can take care of those tasks for you?

A traditional CMS is a big, complex piece of software, because it needs to take into account the many different types of organizations of Web sites. Should anyone on your staff be able to create new articles or only reporters? Which editors should be allowed to post items to the Web? What sorts of headers and footers do you need? What sort of search mechanism do you need? The answers are almost endless, which explains why CMS software can be complex to install and administer.

If you want to publish articles on a regular basis but don't want the administrative overhead associated with a full-fledged CMS, you might want to consider a Weblog. Weblogs, also known by the shortened name blogs, began in the mid-1990s as personal journals, on-line diaries that allow an individual to write and post articles quickly and easily. And although blogs vary considerably in style, their format tends to be fairly uniform, which reduces the complexity of the software, making it easier to configure and administer.

This month, we take an initial look at open-source Weblog software, as well as the standards that have become increasingly prevalent in the blogging community. Along the way, we look at COREBlog, a Zope-based tool that makes it fairly easy to create and administer a Weblog.



Welcome to Altneuland

Now Reuven has a blog, too

[Writing] [Israel and Zionism] [United States]

I was born, raised, and educated in the United States. At the age of 25, I "made aliyah," moving to Haifa, and then moving again to Modi'in. (Modi'in is a new city halfway between Jerusalem and Tel Aviv.) In August 2003, I returned to the United States to begin graduate studies in Learning Sciences at Northwestern University. This Weblog is not only an attempt to reduce the torrent of e-mail I've received from people asking what (and how) I'm doing, but also to pontificate on a number of subjects.

The name of the Weblog, Altneuland ("Old-New Land") is borrowed from the book of the same name, written by Theodor Herzl more than 100 years ago. Herzl wrote Altneuland as a utopian vision of what the future Jewish state might look like, years before such a state might exist. Altneuland, as you might expect, is both brilliantly insightful and laughably off the mark. But it's still worth reading by anyone interested in Israel and Zionism.

I decided to call my Weblog "Altneuland" because (a) it wasn't used by anyone else, (b) it reflects my Zionist thinking, and (c) after eight years in Israel, the United States definitely feels like an old-new land to me. It's very familiar, and yet very strange and different from what I'm used to.

Posted by **reuven** at 16:25 | Comments (0) | Trackbacks (0)

Comments

There is no comment.

Trackbacks

Please send trackback to:http://altneuland.lerner.co.il//2/tbping

There is no trackback.

Figure 1. The author's blog uses the Zope-based COREBlog package.

To be honest, I have some personal interest in finding a good blogging package. Having read a number of Weblogs over the last few years, I've decided it's time to try blogging for myself. The results of my search should be available by the time you read this at blog.lerner.co.il (Figure 1).

What Is a Weblog?

Weblogs come in all shapes and sizes, reflecting their authors' interests and styles. That said, a number of characteristics are common to most Weblogs:

- Order: postings are displayed in reverse-chronological order, with the most recent article displayed at the top of the page. The Weblog's home page typically shows only the last few days of postings, with the rest available through an archive feature.

- Comments: readers of the Weblog are invited to submit comments, often posted immediately following the article in question. In this way, Weblogs are similar to Web-based forums, except only the blogger is allowed to begin a discussion topic.

- One author: typically, only one person participates in a Weblog. Some Weblogs are written by multiple authors, but this is relatively rare. A Weblog's comments, as described above, are written by people other than the main author.

- Syndication: the contents of a Weblog generally are made available using an XML format known as RSS, which stands for really simple syndication. This makes it possible to retrieve, analyze and collect a number of different Weblogs, creating something akin to a personal newspaper.

- Trackback: introduced by the proprietary package Movable Type, trackbacks make it possible for Weblogs to keep track of links pointing to one another.

- Web-based editing: because Weblogs exist only on the Web, their interfaces expect that you want to write and edit postings using your Web browser. Administering a Weblog in this way makes a great deal of sense, but writing and editing in a Web browser can be an unpleasant experience. Many Weblog packages offer desktop applications that ease this pain, allowing users to write using a word processor or simple text editor and upload the final product when they are done. Open-source systems are available that can make such off-line editing possible, but they involve some configuration. Therefore, you probably are going to need to get used to writing inside of a text-area widget or learn how to customize Mozilla so you can use a friendlier editor. You also can look into the open-source Epoz Project, epoz.sf.net, which provides a cross-browser, JavaScript-based editing system.

The above list is not comprehensive. Plenty of Weblogs lack comments, syndication or trackback. But just as English-language newspapers evolved to have a common set of style rules for headlines, captions and story ranking, so too have

Weblogs evolved to have a common set of expectations. And the competition for features is rather fierce: when one Weblog package adds a useful new feature, others usually implement it within a short period of time.

The above features would be easy and straightforward for an experienced Web/database programmer to implement in a high-level language, such as Perl or Python. If you use a relational database, say PostgreSQL, to store the articles, you no longer have to worry about ordering or file storage, so you can concentrate on output. And indeed, some prominent bloggers, such as Tim Bray, author of the excellent Ongoing Weblog at www.tbray.org/ongoing, have created their own Weblog software.

As much as I enjoy writing new programs, I dislike reinventing the wheel. And given the plethora of good, existing solutions for creating a Weblog, including several that allow me to write plugins that extend their functionality, I decided to use something that already exists, extending and modifying it as necessary using an established API.

Moreover, some of the nicer Weblog features, such as comments and trackback, can be tricky to implement. They aren't necessarily hard to work with, per se, but given that it seems 100,000 different mechanisms for commenting on Weblog postings already are out there, I would like to avoid creating number 100,001.

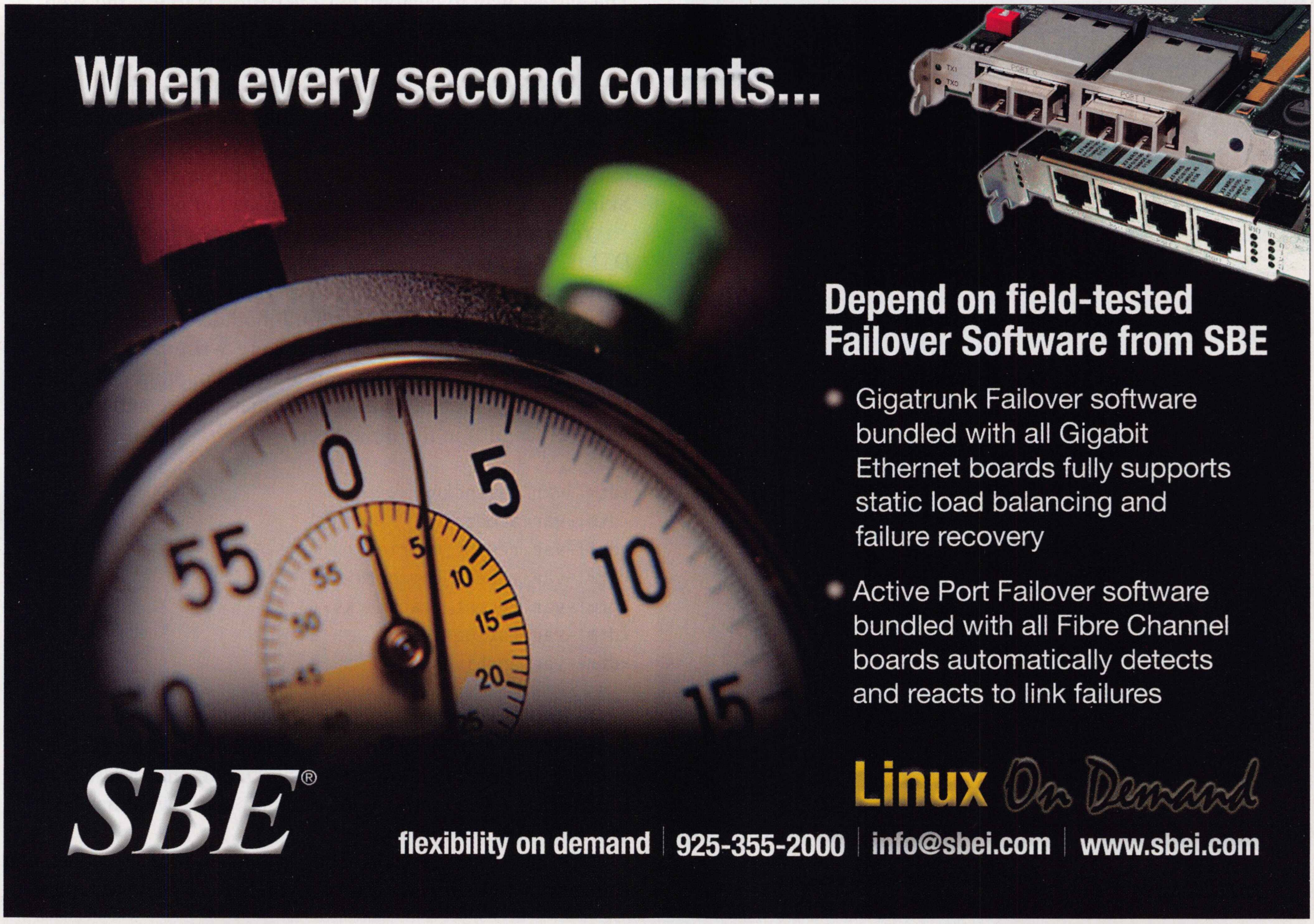
I should note that if you are interested in creating your own Weblog, there is an alternative that allows you to avoid writing or configuring any software at all—namely, using one of the many free Weblog hosting sites on the Internet. These might be a perfectly adequate solution for most people, but I still would like to have some control over the software that I use. Moreover, I would like to integrate my Weblog into the rest of my site and domain, meaning that I need to install it on my own system.

COREBlog

If you are looking for the open-source Weblog package with the simplest installation, and if you already are familiar with the Zope application server, you might want to consider COREBlog. COREBlog, written by Atsushi Shibata, is an actively developed Zope product, or plugin module.

Zope, as I have mentioned in previous installments of this column, is an open-source application server written and distributed by Zope Corporation. Zope is written largely in Python and is object-oriented, using an object database (ZODB) to store most of its core information. Zope development is quite different from other languages and application servers, and adjusting to its mindset can take some time and effort. But it also is quite flexible, making it easy for developers to add their own modules (products) to the system.

COREBlog comes as a standard tarball, which must be opened in the lib/python/Products directory under your Zope root directory. As of this writing, the latest version of COREBlog is 0.53b, which arrives on your system as a file


When every second counts...

Depend on field-tested Failover Software from SBE

- Gigatrunk Failover software bundled with all Gigabit Ethernet boards fully supports static load balancing and failure recovery
- Active Port Failover software bundled with all Fibre Channel boards automatically detects and reacts to link failures

Linux On Demand

SBE®

flexibility on demand | 925-355-2000 | info@sbei.com | www.sbei.com

named COREBlog053b.tgz system when you download it from www.coreblog.org. The following instructions assume that the environment variable ZOPE is set to Zope's root installation directory (/usr/local/zope on my system) and that the tarfile for COREBlog is in /tmp:

```
# cd $ZOPE/lib/python/Products
# tar zxvf /tmp/COREBlog053b.tgz
# chown -R zope.zope # Or appropriate owner/group
```

Restart Zope, either manually or from the Web interface in its control panel, and COREBlog automatically is added as an available product.

To use COREBlog, you need to create an instance of the product from the Add menu in the upper-right corner of the screen. Point your Web browser to the /manage URL on your site (for example, www.example.com/manage), and select COREBlog. You are asked to provide an ID (a unique name to appear in the URL), as well as a title (which appears as the name of the Weblog) and a character encoding (which defaults to ASCII and which I normally change to UTF-8, for full Unicode support).

At this point, the Weblog is almost ready to be unveiled to the world. We can view the Weblog using a URL that ends with the ID we assigned to it (for example, /atf), or we can administer it by appending the /manage path to that name (in our case, /atf/manage).

GUI Builders for Professionals

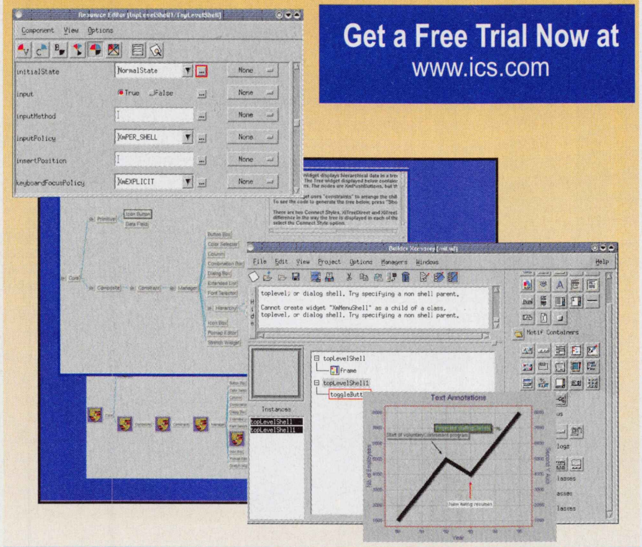
BX PRO

- Visual Layout of World-Class Interfaces
- Round-Trip Editing of Non-UI Code
- Full Support of All New Motif Widgets
- Large Project Support
- Object-Oriented C/C++ Development
- Easy to Install and Configure
- Royalty-Free Distribution of Developed Code
- Thorough Documentation
- Expert Technical Support Included

Get a Free Trial Now at
www.ics.com

ICS
Integrated Computer Solutions Incorporated

www.ics.com 800-704-6824 info@ics.com
201 Broadway Cambridge, MA 02139
The User Interface Company™



Because every posting in COREBlog must be associated with at least one category, we must create at least one category before we can begin to post. Indeed, a warning in red tells us we must add a category before continuing. A Categories tab resides on our blog management screen; click on it, and you are invited to add a new category. We also can rename categories and see how many postings are associated with each category.

Once we have created a category, we can click on the Entries tab and begin to enter new postings. Each entry consists of a minimum of a title, body and category. Other items are either optional or defaults to reasonable values, such as the current date and time. When you have finished posting an item to the blog, you can preview it (which I highly recommend) with the Preview button or publish it right away with the Add button. Once you have published the story, it is visible to the entire world. Anyone reading your Weblog can see the new posting at the top of the page.

But someone doesn't have to look at the top of the page in order to notice the new posting. The recent entries area of the sidebar contains links to each of the most recent postings. The calendar has a hyperlink to all of the postings made on a particular day. And clicking on the topic displays all entries on that topic. This works only for topics, though, not for subtopics.

Customization

The easiest way to customize COREBlog is to use the Web-based properties editing tool, labeled Settings. The Settings tab allows you to write and edit the content and behavior of the various pages on the site. For example, you can write a bit about yourself or your Weblog or indicate that comments are to be moderated by default. If you are interested in changing the fonts and colors in which your blog is displayed, click on the Skins tab to gain access to such information.

You can customize COREBlog further by modifying the DTML pages used to display output. In particular, you can change the sidebar (along the right side of the page) by modifying files in COREBlog/dtml/modules (under $ZOPE/lib/python/Products). These pages are written in DTML, Zope's original server-side templating language. The file index_html.dtml does nothing more than invoke and use each of the other files in the directory:

```
<dtml-var calendar>
<dtml-var about>
<dtml-var recent_entries>
<dtml-var recent_comments>
<dtml-var recent_trackbacks>
<dtml-var categories>
<dtml-var archives>
```

In order to change the order of modules or which modules appear at all, simply modify index_html.dtml. You also can create your own modules, write new DTML files and use the COREBlog API to retrieve information about what has been said and done on the blog.

Of course, because COREBlog is an open-source product, you can view and modify any part of it you like. To be honest, there isn't much to modify; most of the features in COREBlog probably are ones that you want to include in your Weblog. Given that the project is under active development, new features should be added soon.

Syndication

When you become an established blogging pundit, people will turn to your site several times a day, looking for your latest wisdom and pointers to interesting links. People who follow more than one blog or who want to integrate Weblogs with some news sources use an aggregator program to collect content from various sites. Aggregators do not look at the HTML directly; rather, they retrieve the syndication feed as distributed in RSS.

COREBlog makes it easy to syndicate your site. Indeed, you don't need to do anything in order to syndicate postings on your Weblog. COREBlog takes care of this for you automatically, providing both the popular RDF and RSS formats. There is nothing to configure in order for syndication to work; COREBlog comes with working syndication out of the box.

Conclusion

Content management systems are wonderful for organizations that need them. But if you are running a Web site by yourself, a full-fledged CMS probably is overkill. Many individuals have begun to publish Weblogs in the last few years, and although Weblog software still is in a relatively early stage, such products do make it easy for an individual to get started publishing articles on a regular basis, in an easy-to-understand, standard format. COREBlog is one open-source application for creating and managing Weblogs, and it seems to do an excellent job of offering the basics. It also provides a plugin architecture so new modules can be added to the system without having to restart it from scratch. Next time, we continue on our tour of open-source Weblog software, comparing some other packages to COREBlog.

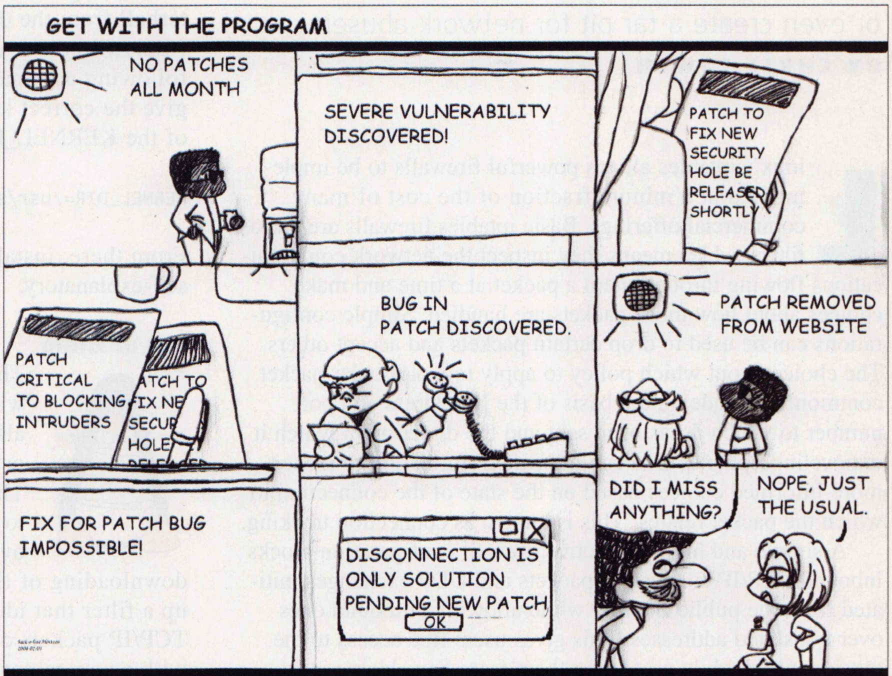
Reuven M. Lerner, a longtime consultant in Web/database programming, is now a graduate student in Learning Sciences at Northwestern University in Evanston, Illinois. You can reach him at reuven@lerner.co.il.

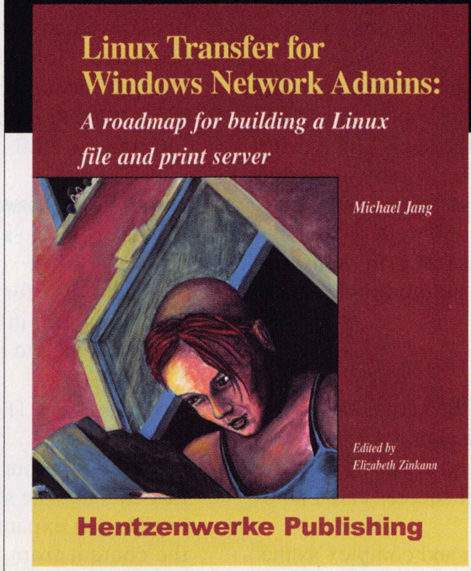
RESOURCES

COREBlog is available from **www.coreblog.org**. Zope is available from **www.zope.org**, and Python is available from **www.python.org**.

A good introduction to trackback is written by the authors of Movable Type; see **www.movabletype.org/trackback/beginners** or **www.cruftbox.com/cruft/docs/trackback.html**.

For information about RSS and syndication, take a look at **www.xml.com/pub/a/2002/12/18/dive-into-xml.html** or **www.webreference.com/authoring/languages/xml/rss/intro/2.html**.


GET WITH THE PROGRAM
NO PATCHES ALL MONTH
SEVERE VULNERABILITY DISCOVERED!
PATCH TO FIX NEW SECURITY HOLE BE RELEASED SHORTLY
PATCH CRITICAL TO BLOCKING INTRUDERS
ATCH TO FIX NE SECUR HOLE RELEASED
BUG IN PATCH DISCOVERED.
PATCH REMOVED FROM WEBSITE
FIX FOR PATCH BUG IMPOSSIBLE!
DISCONNECTION IS ONLY SOLUTION PENDING NEW PATCH
OK
DID I MISS ANYTHING?
NOPE, JUST THE USUAL.


Linux Transfer for Windows Network Admins:
A roadmap for building a Linux file and print server
Michael Jang
Edited by Elizabeth Zinkann
Hentzenwerke Publishing

Need to move your file server to Linux in a hurry?

"Linux Transfer for Windows Network Admins" will show you how to convert your existing Windows PDC and file/print server to Linux, using your knowledge and skills as a Windows admin to get the job done quickly and easily.

Download free sample chapters and order online at www.hentzenwerke.com

Also available at selected bookstores nationwide

ISBN 1-930919-46-8 298 pages $49.95

Hentzenwerke

The Hidden Treasures of iptables

With these powerful add-ons for iptables you can match strings or port ranges in iptables rules or even create a tar pit for network abusers.

BY CHRIS LOWTH

Linux's iptables allows powerful firewalls to be implemented at a minute fraction of the cost of many commercial offerings. Basic iptables firewalls are packet filters, which means they inspect the network communications flowing through them a packet at a time and make choices about how those packets are handled. Simple configurations can be used to drop certain packets and accept others. The choice about which policy to apply to a particular packet commonly is made on the basis of the IP address and port number to which it has been sent and the direction in which it is traveling. iptables also can use state information to make more-informed choices based on the state of the connection to which the packet relates. This is known as connection tracking.

A simple and highly effective firewall configuration blocks inbound TCP/IP connection packets and UDP exchanges initiated from the public Internet while allowing outbound ones over translated addresses. This gives users free access to the outside world while protecting them from unwelcome intrusions. Such configurations are a bit simplistic and may need additional filters to be truly useful, but the basic concept is straightforward.

iptables has a lot more to offer than these simple packet-filtering criteria. Some of the extras are fairly well known and even may make their way into some off-the-shelf Linux distributions, but some lesser-known features are worthy of investigation. These are the hidden treasures I intend to point you toward in this article. It would take a book to describe all the possible features and options associated with them, so all I do here is flag their existence and put you on the path of exploration.

Introducing the POM

Netfilter has two groups of components, the kernel and user-mode pieces. The user-mode group consists of the iptables and related utilities, libraries, manual pages and scripts. The kernel components are patches to existing kernel sources and a number of extra modules.

Applying patches to a system as large and complex as the Linux kernel can be a daunting task to the uninitiated, and the road is littered with traps and potential blind turns. A bad or incompatible patch readily can produce a kernel that doesn't compile, or worse, doesn't boot. The Netfilter team has sought to resolve these difficulties by providing us with a robot guide, POM, or Patch-o-matic. POM is a collection of patches and a script for applying them to your kernel, and it's a joy even for a relative novice to use.

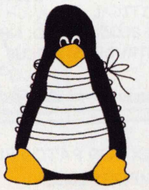
The kernel patches included with POM are classified into a number of groups according to their history and quality. Some of them are base patches needed in every iptables/Netfilter installation. Others are optional or experimental extras that provide interesting features, some of which I describe in this article. These are the promised hidden treasures, what the POM documentation describes as "Maybe broken, Maybe cool extensions."

Running POM is simple; download the latest Patch-o-matic tarball from the directory /pub/patch-o-matic on ftp.netfilter.org, restore it on your system and run the following command while logged in as root. Make sure to give the correct kernel source directory name as the value of the KERNEL_DIR parameter:

```
KERNEL_DIR=/usr/src/linux-2.4 ./runme extra
```

From there, installation is interactive and more or less self-explanatory.

Bits of String



The string module probably is the most widely used extra from the POM trove. It allows packets to be matched against strings occurring anywhere in their data payload. This module has all sorts of uses but needs to be applied carefully so as not to be overzealous. One possible use is to block the downloading of ELF executables from the Web. We can set up a filter that identifies Web return traffic by looking for TCP/IP packets coming from the Internet-facing interface with a source port of 80. If we know that an ELF file starts with hex character 7f followed by the letters ELF (which it does), we can use the string match to search for this sequence. Non-ASCII characters can be embedded in the string by using the pipe symbol to enclose them, so we use `|7F|ELF`. Assuming that the Internet-facing network interface is eth0, the command is:

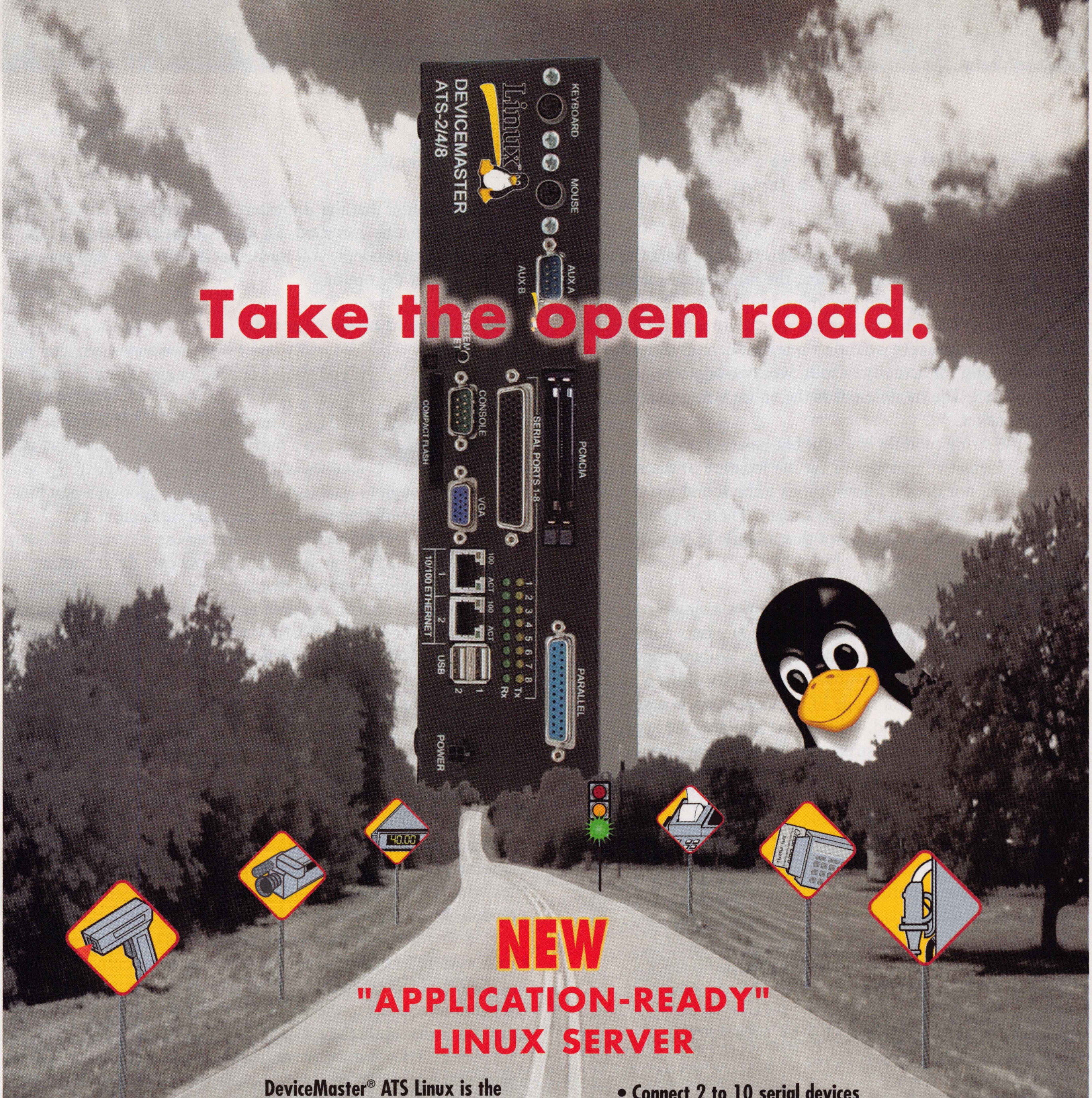
```
iptables -A FORWARD -i eth0 -p tcp --sport 80 \
    -m string --string´|7F|ELF´ -j DROP
```

The syntax for embedding hex characters into the string was introduced in iptables 1.2.8. If you are using an earlier version, you need to resort to trickery. For example:

```
--string "`dd if=/bin/ls bs=4 count=1 2>/dev/null`"
```

takes the first four characters of /bin/ls, which is an ELF file that contains the string we want.

We can expand this example by declaring that we trust the content from 192.168.0.5 and, therefore, don't want to apply the filter to that server. This is done easily by adding an inverted match on the IP address, like this:


Take the open road.
Linux
DEVICEMASTER
ATS-2/4/8
KEYBOARD
MOUSE
AUX A
AUX B
SYSTEM RESET
CONSOLE
COMPACT FLASH
SERIAL PORTS 1-8
PCMCIA
VGA
10/100 ETHERNET
1
2
100
ACT
Tx
Rx
PARALLEL
USB
POWER
40.00

NEW
"APPLICATION-READY" LINUX SERVER

DeviceMaster® ATS Linux is the ideal platform for your next data acquisition, device control, or remote software server. This embedded microcomputer is a fully integrated, fully tested, network-ready Linux server appliance. Add your software to reduce the time-to-market and eliminate the cost and frustration of building your own server. Highlights include:

- 300Mhz processor, 128MB SDRAM, 512MB Compact Flash, Debian Linux 3.0 operating system
- Headless, solid-state design (no fan/disk)
- Connect 2 to 10 serial devices
- Dual 10/100Base-T Ethernet ports
- Add peripherals via USB or PCMCIA

DeviceMaster ATS Linux is available now from Comtrol distribution partners worldwide. Visit us at **www.comtrol.com/ats-linux** to learn more about how you can get the jump on your competition!

For more information, call Comtrol at 800-926-6876.

COMTROL®
The Enterprise Connection
www.comtrol.com

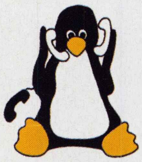
© 2004 Comtrol Corp. All trademarks used herein are the property of their respective trademark holders.

```
iptables -A FORWARD -i eth0 -p tcp ! \
   -s 192.168.0.5 --sport 80 -m string \
   --string ´|7F|ELF´ -j DROP
```

This example has a couple of problems that highlight the issues with the string match module. First, the rule matches any packet that contains this sequence anywhere in the data, not only at the start of the file. This means the rule could match false positives and block packets we didn't intend. Second, if the string we are looking for actually is split over two adjacent packets, it isn't matched. The module needs the entire string to appear in a single packet.

So, the string module is useful but basic. It doesn't allow for case-insensitive matches or for the location of the string to be specified, nor does it allow strings to be found when split over multiple packets in the data stream. There is plenty of scope for an extended version of this module to be written.

Fewer Rules with mport



The mport extension allows a single rule to specify a number of port numbers and ranges using an extended syntax. Without mport, the iptables command can specify either a single port or a range of adjacent ports in a single command. With mport in place, the syntax allows more complex constructs. For example, we could permit X terminals, Web and mail with a single command, like this:

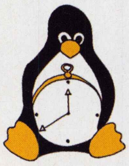
```
iptables -A INPUT -p tcp -m mport \
   --dports 80,110,21,6000:6003 -j ACCEPT
```

Without using mport, this would have to be specified using four separate commands:

```
iptables -A INPUT -p tcp --dports 80 -j ACCEPT
iptables -A INPUT -p tcp --dports 110 -j ACCEPT
iptables -A INPUT -p tcp --dports 21 -j ACCEPT
iptables -A INPUT -p tcp --dports 6000:6003 \
      -j ACCEPT
```

Using a single rule in place of four offers a potential performance advantage because packets passing through the system require less processing. It also makes the maintenance of the rules files easier because services requiring identical processing can be grouped together easily. As you probably guessed, mport is short for multiple ports.

Time-Based Rules

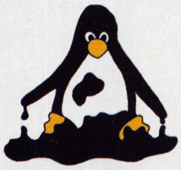


The time module allows rules to introduce the time of day and the day of the week into matching logic. Example uses would be to allow access to personal Web sites only during the lunch hour or to divert Web traffic to a secondary server during routine maintenance periods. The following example renders the Web service inaccessible between the hours of 4 and 6:30AM on Fridays, presumably for system maintenance:

```
iptables -A INPUT -p tcp -d 80 -m time \
   --timestart 04:00 --timestop 06:30 --days Fri \
   --syn -j REJECT
```

It is worth noting that the -timestart, -timestop and -days options all must be specified. So if you want a rule that is not day-of-week dependent, you must specify all seven day names; you can't omit the option.

Getting Bogged Down—Tar Pits



You really don't want to wander into a tar pit if you value your life or appreciate changes of scenery. They are nature's equivalent of fly paper; come too close and you won't leave in a hurry. The TARPIT component of iptables is the networking equivalent: if you are unwise enough to establish a TCP/IP connection to a port that is a tar pit, you will find it hard to close the connection and release the used system resources for future use.

To achieve this tar pit state, iptables accepts the incoming TCP/IP connection and then switches to a zero-byte window. This forces the attacker's system to stop sending data, rather like the effect of pressing Ctrl-S on a terminal. Any attempts by the attacker to close the connection are ignored, so the connection remains active and typically times out after only 12–24 minutes. This consumes resources on the attacker's system but not the Linux server or firewall running the tar pit. You could use the following iptables command to pass packets to the pit:

```
iptables -A INPUT -p tcp -m tcp -dport 80 -j TARPIT
```

You probably don't want to use conntrack and TARPIT on the same system, particularly if you anticipate catching a lot of flies with this particular brand of fly paper. Each stuck connection consumes conntrack resources.

One way to confuse potential attackers is to make your Linux system look like a Microsoft Windows machine by causing the netbios ports to respond to port scans. Then pass any connection requests to the tar pit. This has the effect of wasting attackers' time while they sense a possible opening and try to gain access. They will be frustrated by long timeouts and an apparently buggy target. Rules such as the following produce this result:

```
iptables -A INPUT -p tcp -m tcp -m mport \
   --dports 135,139,1025 -j TARPIT
```

Another possibility is to TARPIT all ports except the ones you genuinely want to use. This again leads outsiders to see every port as open and waste time attempting to gain access. Moreover, a configuration like this prevents tcpdump from correctly determining the operating system running on the server. In this example, we allow Web and e-mail traffic and bog down everything else:

```
iptables -A INPUT -p tcp -m tcp --dport 80 -j ACCEPT
iptables -A INPUT -p tcp -m tcp --dport 25 -j ACCEPT
iptables -A INPUT -p tcp -m tcp -j TARPIT
```

You can find an interesting real-life story of how TARPIT and string helped one particular system administrator (not me) at www.spinics.net/lists/netfilter/msg17583.html.

Get a closer look inside your network

Cyclades-nQuirer
cyclades
nQuirer
Network Monitoring
LINUX INSIDE

Introducing the Cyclades-nQuirer with NetFlow™ support

The complex networks of today carry all different types of traffic, from customer information to viruses that can bring your network to a halt. Having visibility over your network traffic is essential for efficient network management and planning.

With the Cyclades-nQuirer, you can gather traffic data and convert it to NetFlow™ format, which allows you to consolidate and analyze your network in a higher-level format and achieve faster problem solving and increase network uptime. All this in compact form factor and at a fraction of the cost of comparable solutions.

With the Cyclades-nQuirer you will be able to look inside your network and understand it like you've never had before.

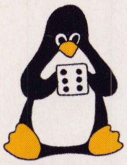
LINUX JOURNAL READERS' CHOICE 2003
"Best Hardware for Linux since 1995"

www.cyclades.com/ljb
1.888.cyclades ▪ sales@cyclades.com

cyclades
Everywhere with Linux

©2004 Cyclades Corporation. All rights reserved. All other trademarks and product images are property of their respective owners. Product information subject to change without notice.

Randomizing



The random match module matches packets based on nothing more than a random choice. You can tune the logic by setting the probability that a packet is matched anywhere between 0% and 100% of the time. Example applications include simulating a faulty connection or server or distributing load across multiple mirrored Web servers. The example below distributes Web traffic among three servers. The first rule sends 33% of the connections to the server at 192.168.0.100. The next 33% is sent to 192.168.0.101 and the last third catches the remainder and passes them to 192.168.0.102:

```
iptables -t nat -A PREROUTING -i eth0 -p tcp \
   --dport 80 --syn -m random --average 33 \
   -j DNAT --to-destination 192.168.0.100:80

iptables -t nat -A PREROUTING -i eth0 -p tcp \
   --dport 80 --syn -m random --average 50 \
   -j DNAT --to-destination 192.168.0.101:80

iptables -t nat -A PREROUTING -i eth0 -p tcp \
   --dport 80 --syn -j DNAT \
   --to-destination 192.168.0.102:80
```

As before, this assumes that eth0 is the Internet-facing interface.

A Lot More Where They Came From

Dozens of treasures can be dug up and enjoyed. I have described a small handful here, but there are plenty more. Simply running the `runme` script and reading the patch descriptions as they are displayed is one way of getting an idea of what is available. Here are a few more examples of what you can find:

- Connection tracking for RSH, MMS (media streaming), PPTP, Quake, RPC and Talk.
- Extended support for configuration and status information access through the /proc filesystem.
- Extended support for IPv6 features.
- Manipulation of options, TTL and more in IP packets.
- Finer control over NATed connections.
- Control over limits on quota and bandwidth usage.
- Anti-OS fingerprinting logic and port-scan detection.
- Connection marking (and mark testing).

Sources of Wisdom

The patches added with POM don't add their descriptions to the iptables man page, so we need to turn elsewhere for documentation. The basic syntax used to invoke these extensions can be displayed using the iptables built-in help facility. For example, `iptables -m random -help` gives the usual help message but with the random module's parameters displayed at the end. The same technique can be applied to the other modules.

You also can refer to the module help files held in the Patch-o-matic directory structure. The file for random is base/random.patch.help. Similar files can be found for the other patches.

Finally, make use of the Netfilter Web site, www.netfilter.org/patch-o-matic, which contains a description of each of the POM patches.

Installing New iptables Modules

The majority of iptables extensions have two parts, a patch to the Linux kernel and a configuration helper library for use by the iptables user-space program. The detailed procedure for adding a POM module to the kernel and the iptables tools is outlined at www.lowth.com/howto/add-iptables-modules.php. In summary, the steps we need to take are bring your system up to date; download the latest Patch-o-matic sources; patch the kernel using the runme script; recompile and install the patched kernel; and recompile and install the iptables software.

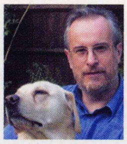
Conclusion

We have seen that Linux's Netfilter provides an excellent set of features for building effective firewalls, but not all of these features are installed by default on many Linux distributions. The Patch-o-matic software allows administrators to extend the base functionalities of their firewalls through an automated approach to patching the Linux kernel.

To finish, take this thought with you: we have seen that iptables/Netfilter has a number of exciting possibilities hidden away from initial inspection. The chances are high that the same is true for other packages. This is part of the joy of open-source software; nothing is truly hidden. Everything that exists is waiting there for the skilled seeker to find.

Acknowledgement

To Jane Lowth for drawing the figures of Tux.



Chris Lowth works as a consultant for Intercai Mondiale (www.intercai.co.uk), a UK-based telecommunications, IT and business consultancy. He designs security software and network management (OSS) solutions and attempts to play the guitar. Chris can be contacted at chris@lowth.com.

TECH TIP

During periods of e-mail worm activity, like now (news.google.com/news?q=new%20email%20virus), put this in your ~/.spamassassin/user_prefs to keep your mailbox worm-free:

```
score MICROSOFT_EXECUTABLE 24
```

You can comment out the line if you ever want to receive Microsoft executables at some time in the future.

PENGUIN COMPUTING
Altus 1000E
POWER
HDD
PENGUIN COMPUTING
Relion 130
POWER
HDD
PENGUIN COMPUTING
Relion 230
PWR
LAN
HDD
HDD
PENGUIN COMPUTING
Altus 3200
POWER HDD
PENGUIN COMPUTING
Relion 430

The World's Most Reliable Linux Systems

clusters

servers

workstations

Scyld Beowulf™

www.penguincomputing.com

1-888-penguin

© 2003 Penguin Computing. All rights reserved. Penguin Computing and the PC logo are registered trademarks of Penguin. Linux is a registered trademark of Linus Torvalds. Scyld Beowulf is a trademark of Scyld Computing Corp. Photo © 2003 Gary Wagner.

François, Can You Keep a Secret?

Learn the fundamentals of public key encryption and how to plug the GNU Privacy Guard in to your existing mailer. **BY MARCEL GAGNÉ**

It really is a shame François, but we simply will have to make sure it is available for next time. Everything else looks perfect, *mon ami*; all the workstations booted up to a login. Wonderful! Ah, I see that our guests are here! François, head down to the east wing of the wine cellar. There is a small cache of 1999 Côte-Rôtie next to that old sealed-in door. Yes, the door we never were able to open. *Vite, François!* I promise you there is nothing frightening down there.

Please sit, *mes amis*. We have an almost perfect menu for you today, but sadly it's missing one item. Still, the Côte-Rôtie Rhône, red that is both sexy and mysterious, should help take away from its absence. I had so wanted to prepare my famous *Crème Linuxaise* for you today, but there were problems. You see, the Linuxaise is an old and secret family recipe, and I could not risk it falling into the wrong hands. Nor could I risk sending it by e-mail for fear of it being intercepted or read by a network sniffer; otherwise, François could have prepared it in advance. It is that secret!

Next time, however, will not be a problem. I am setting up all users at the restaurant with GnuPG and public key encryption so that sensitive communications can be sent without fear. GnuPG is the GNU Privacy Guard, a program that makes it possible to encrypt messages and data in general. It is a patent-free, open-source replacement for PGP (Pretty Good Privacy). A number of Linux e-mail packages allow you to send and receive encrypted e-mails using GnuPG, and this is what I'd like to show you today. Most modern distributions have GnuPG included, so if you don't have it on the system, check your CD first. You also can find the latest version at www.gnupg.org, but first, a little background.

Now, where was I? Ah, yes. Historically, all encryption methods worked on the premise of a shared key file. You would give the person with whom you wanted to communicate the same key by which the message was encoded. Think secret decoder ring and you aren't far off the mark. The catch is that anyone intercepting the key then could decipher all your messages. With GnuPG, messages are encrypted with two keys, one being your private key, which you guard jealously and never hand out to anyone. When I encode a message, I do so by combining my private key with a public key, not my public key, but one supplied to me by the person with whom I want to communicate, François, for instance. Both keys are required for the encryption/decryption process, and anyone having one-half of the key pair has nothing, which is why you never hand out your private key. To get in on this top-secret action, you need to create a key pair, containing both your private and jealously guarded key and your public key, the one you hand out to all your friends. Here is the command:

```
gpg --gen-key
```

What follows is a small question-and-answer session. The first question has to do with the encryption algorithm, or cipher. The default is DSA and ElGamal. Accept the default. When asked about key length, you can choose from 768, 1,024 and 2,048 bytes. Because the DSA standard is 1,024, choose that for now. Then, you are asked for the expiration date of your key, the default being no expiration date. You can, however, define days, weeks, months and even years. For now, choose the default and confirm your choice. Finally, you are asked to supply the name of the key user, the e-mail address and a comment.

It's all over but the passphrase, which is your final step. Make sure you choose something secure but also something that you can remember. When you have entered your phrase, the GnuPG program generates your secure key pair. After the command completes its work, you can verify the result by looking in the .gnupg directory in your home directory:

```
$ cd /home/marcel/.gnupg
$ ls
gpg.conf     pubring.gpg  random_seed
secring.gpg  trustdb.gpg
```

The gpg.conf file contains a list of default options for the gpg command and makes for good reading. The pubring.gpg and secring.gpg files are particularly important. Make backups of these files immediately and store them somewhere other than your computer and never lose them—the secring.gpg file contains your personal secret key. The last file, trustdb.gpg, is your database of trust. It defines the level of trust that you assign to the public keys you collect.

Users need to exchange keys in order to exchange encrypted information. As you might guess, the option to export a key is --export, but you may want to include the -a option as well to confine the output to ASCII format:

```
gpg -a --export 3E2FCF7D > marcelkey.asc
```

The resulting file is a simple ASCII text file; how you get it to the other person is up to you. There are key servers where you can upload your public keys so that anyone can download them (which makes it easier for large-scale distribution), key-signing parties where groups get together and exchange public keys or e-mail attachments. The recipient then could import the key like this:

```
gpg --import marcelkey.asc
```

At any point, you can choose to see the keys in your key ring with:

```
gpg --list-keys
```

www.nexcom.com

Blade Server with Dual XEON™ Platform

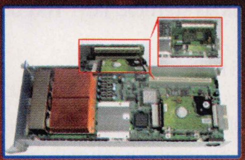
4U Enclosure, Xeon™ 3.06GHz, PCI-X Expansion

NexBlade HS416

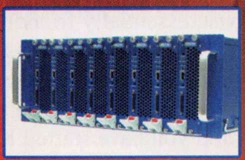
SPECIAL SAMPLE OFFER

Item	Model	Q'ty
1	HS 416-BMS chassis w/350W PSU x 2 and 12 ports Ethernet Switch	1
2	HDB42732 server Blade w/Dual Xeon 2.8GHz, 512MB DDR SDRAM and 40GB HDD(2.5")	1
3	HDB42732I3 server Blade w/Dual Xeon 2.8GHz, 512MB DDR SDRAM, 40GB HDD(2.5") and one PCI-X expansion slot	1

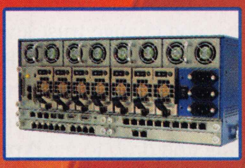
US$7,800



- **OFF-THE -SHELF EXPANSION**
  ONE PCI-X
  EXPANSION SLOT
  ON EACH SERVER BLADE



- **COMPACT AND POWERFUL**
  8 SERVER BLADES WITH DUAL
  XEON 3.06 IN 4U ENCLOSURE



- **VERSATILE NETWORK CONNECTIVITY**
  SWITCH AND/OR FEEDTHRU BLADE FOR
  GIGABIT OR FAST ETHERNET

NEXCOM®
The Digital Infrastructure

NEX COMputers Inc.
46706 Fremont Blvd., Fremont, CA 94538,U.S.A.
Tel: 510-656-2248 Fax:510-656-2158
e-mail: sales@nexcom.com

The results of that command depend on how many keys you have, but the following should give you an idea of what to expect. In particular, take note of the hexadecimal number following the 1024D. That is the key ID to which you will be referring in the future:

```
/home/marcel/.gnupg/pubring.gpg
-------------------------------
pub  1024D/3E2FCF7D 2004-01-07 Marcel Gagné
➥(Writer and Free Thinker at Large) <mggagne@salmar.com>
sub  1024g/B24717BE 2004-01-07

pub  1024D/EE392B87 2004-01-07 Francois
➥(I am but a humble waiter) <francois@salmar.com>
sub  1024g/F4E07040 2004-01-07
```

Before you can start your friend's key to encrypt e-mail messages, you need to sign the key to verify its authenticity. This involves doing two things. If you are absolutely, positively sure of the key's origins, you might skip the first step, which is to get the key's fingerprint:

```
$ gpg --fingerprint francois
pub  1024D/EE392B87 2004-01-07 Francois
➥(I am but a humble waiter) <francois@salmar.com>
Key fingerprint = 8C5B 775C 33F8 E97C 5ADC
➥019D C6C8 4B83 EE39 2B87
sub  1024g/F4E07040 2004-01-07
```

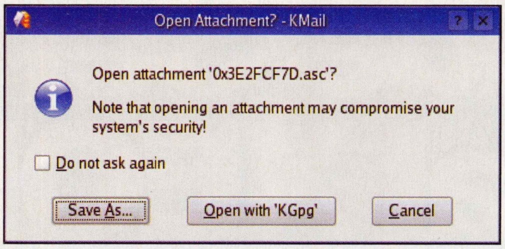
Notice that I used the person's name in the above command, which is part of the key information. If you have more than one person with that name in their key information, specify the key ID instead. To verify this fingerprint, ask your friend to check the fingerprint of his or her personal key in exactly the same way. The final step is to sign the key. You do that with the --edit-key option:

```
gpg --edit-key francois
```

The whole process isn't too complicated. You are asked to confirm that you really, genuinely want to sign this public key; you are asked to confirm again with your passphrase. You need to do this with all the people with whom you want to exchange encrypted e-mail. Once finished, however, let the secret messages flow.

Incidentally, a number of nice, graphical utilities for GnuPG exist, essentially friendly wrappers for the command-line utility. KDE comes with a slick utility called KGpg that integrates nicely with the desktop and the e-mail system. For instance, if someone sends you a public key as an attachment in KMail and you have started KGpg (command name: kgpg), rather than having to save the attachment to a file, drop down to the command line and perform the above steps. You should get a friendly pop-up like the one in Figure 1.

With this tool, which docks in your panel, you can edit keys; add, remove and change trust levels; and do all those things you can do with command-line GnuPG but with a click of your mouse. You even can do key server lookups over the Net and use photo IDs as well. Its integration with other KDE tools means you can drag and drop to encrypt or access GnuPG functions with a click from Konqueror or the clipboard. KGpg is an add-on to KDE 3.1, but with the release of KDE 3.2, it will become part of the standard distribution. Figure 2 shows KGpg in action.

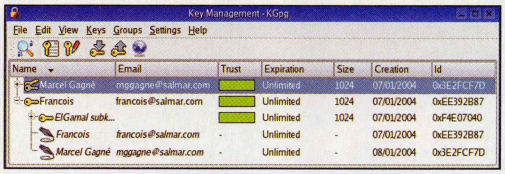


Open attachment '0x3E2FCF7D.asc'?

Note that opening an attachment may compromise your system's security!

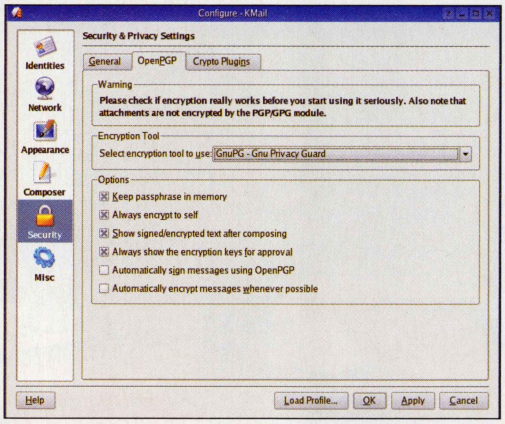
☐ Do not ask again

Figure 1. KMail confirms key imports.



Name	Email	Trust	Expiration	Size	Creation	Id
Marcel Gagné	mggagne@salmar.com	🟩	Unlimited	1024	07/01/2004	0x3E2FCF7D
Francois	francois@salmar.com	🟩	Unlimited	1024	07/01/2004	0xEE392B87
ElGamal subk...		🟩	Unlimited	1024	07/01/2004	0xF4E07040
Francois	francois@salmar.com	-	Unlimited	-	07/01/2004	0xEE392B87
Marcel Gagné	mggagne@salmar.com	-	Unlimited	-	08/01/2004	0x3E2FCF7D

Figure 2. KGpg makes key management easy.



Security & Privacy Settings

Warning

Please check if encryption really works before you start using it seriously. Also note that attachments are not encrypted by the PGP/GPG module.

Encryption Tool

Select encryption tool to use: GnuPG - Gnu Privacy Guard

Options

☒ Keep passphrase in memory
☒ Always encrypt to self
☒ Show signed/encrypted text after composing
☒ Always show the encryption keys for approval
☐ Automatically sign messages using OpenPGP
☐ Automatically encrypt messages whenever possible

Figure 3. Configuring KMail to Use GnuPG Encryption

Another graphical administration tool worth your time is gpa. This is the default GNU Privacy Assistant and the official key ring editor of the GnuPG Project. It also is available from www.gnupg.org.

Now, it's time to send out those encrypted e-mails. Open KMail, click on Settings in the menubar and choose Configure KMail, after which the Configure KMail dialog appears. In the sidebar to the left, find the icon labeled Security, and click it (Figure 3).

The new window to the right has three tabs: General, OpenPGP and Crypto Plugins. We are interested in the OpenPGP tab. On that tab, look at the drop-down box labeled Select encryption tool and choose GnuPG - GNU Privacy Guard from the list. For the time being, don't choose to sign and encrypt all messages automatically. You might, however, want to keep the passphrase in memory; you still are asked the first time an encrypted message is encountered. Click Apply.

Next, click on the Identity tab to the left. Unless you have defined multiple e-mail identities, there should be only one entry. Click Modify, and select the Advanced tab from the resulting dialog. In the middle of that panel (Figure 4) is a space for OpenPGP key. That's your personal private key. Click Change here to open another window from which you can select your private key information.

When finished, click OK to close the Edit Identity window, and click OK again to close the Configure KMail window. In sgi

sgi altix 350
sgi
sgi altix 3000
sgi
sgi

Anything they can crunch, we can crunch better. The Altix® family of servers and superclusters combine the industry leading performance of Intel® Itanium® 2 processors with a production-class Linux® solution. The result is a new class of reliable, cost-effective servers that meet the most demanding requirements. Last year, the Altix® 3000 series took the open standards community by storm, delivering scalability across hundreds of processors and terabytes of memory. Now Altix® 350 brings best-in-class price and best-in-class performance to the mid-range. Whether you need a solution with infinite scalability, or dream of incredible power at an affordable price, there's an Altix system for you. sgi.com/altix

intel inside ITANIUM 2

SILICON GRAPHICS | The Source of Innovation and Discovery™

© 2004 Silicon Graphics, SGI, Altix and the SGI logo are registered trademarks and The Source of Innovation and Discovery is a trademark of Silicon Graphics, Inc. in the United States and/or other countries worldwide. Intel, Itanium, Intel Inside, and the Intel Inside logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. Linux is a registered trademark of Linus Torvalds, used with permission by Silicon Graphics, Inc.

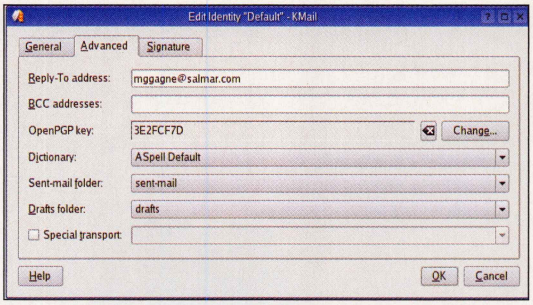
order to reload the keys into KMail properly, you may have to shut down KMail and restart it.

You now are able to send encrypted e-mails, in principle at least. If you already have your friend's public key in your key ring and you have signed the key, nothing is left to do but write your e-mail. Start a new message, and write what you need to write. When you are ready to send, click Options on the message's menubar. Notice two interesting options here related to encryption. One says Sign Message, and the other says Encrypt Message.

Signing a message makes use of your key by attaching an electronic signature (your public key) to your message but not encrypting it. The person receiving the message then has a means of verifying that the message did indeed come from you, should they want to verify it. Because the message is not encrypted, anyone still can read it. This is simply a means to confirm that the message came from the person who claims to have sent it. Encryption goes one step further in that you are using your friend's public key to encrypt the message. In both cases, KMail asks you for your passphrase before sending the message.

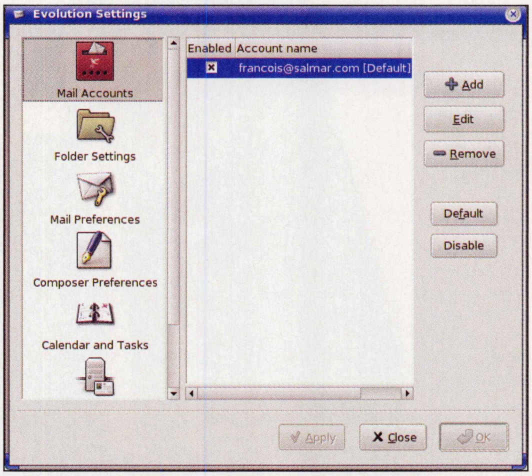
Unfortunately, not all e-mail packages follow the same rules for encryption. That would be too simple, *non*? Let's use the two gentlemen at table 17, Larry and Michael, in a hypothetical scenario. Each needs to send the other sensitive corporate information, so the information must be encrypted. Larry uses Evolution and Michael uses KMail. We already have covered KMail for sending encrypted mail, so let's spend a moment looking at the process in Ximian Evolution.

DesktopLinux.com
. . . join the Desktop Linux revolution!


Edit Identity "Default" - KMail
General
Advanced
Signature
Change...
Help
OK
Cancel

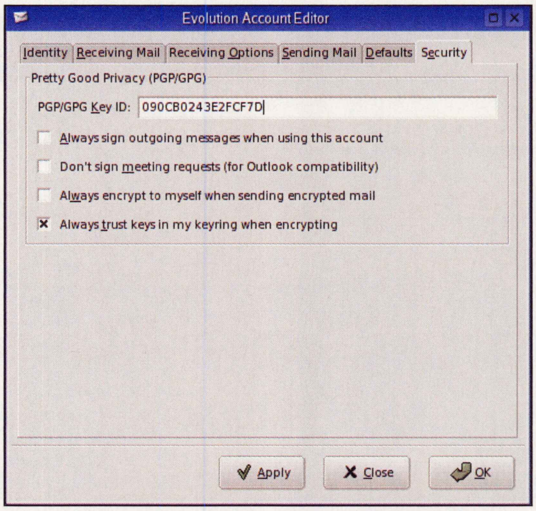
Reply-To address: mggagne@salmar.com
BCC addresses:
OpenPGP key: 3E2FCF7D
Dictionary: ASpell Default
Sent-mail folder: sent-mail
Drafts folder: drafts
☐ Special transport:

Figure 4. Specifying Your Public Key with KMail


Evolution Settings
Mail Accounts
Folder Settings
Mail Preferences
Composer Preferences
Calendar and Tasks
Enabled
Account name
Add
Edit
Remove
Default
Disable
Apply
Close
OK

☒ francois@salmar.com [Default]

Figure 5. Evolution's Mail Account Editor


Evolution Account Editor
Identity
Receiving Mail
Receiving Options
Sending Mail
Defaults
Security
Pretty Good Privacy (PGP/GPG)
Apply
Close
OK

PGP/GPG Key ID: 090CB0243E2FCF7D
☐ Always sign outgoing messages when using this account
☐ Don't sign meeting requests (for Outlook compatibility)
☐ Always encrypt to myself when sending encrypted mail
☒ Always trust keys in my keyring when encrypting

Figure 6. Specifying Your GnuPG Key in Evolution

With Evolution open, click on your inbox. Now, click Tools on the menubar and select Settings. You now are looking at the Evolution Settings window (Figure 5). If the Mail Accounts icon in the left-hand sidebar isn't selected, click to select it now. Most people have one main account. If you have more than one, select the one you want to use with encryption and click the Edit button to the right. The Account dialog appears with several tabs along the top. The one you are interested in is labeled Security. Look for a field labeled PGP/GPG Key ID and enter your key ID there. Click OK to close that dialog, then OK again to get out of the Evolution Settings window.

Assuming that Larry and Michael already have exchanged public keys, Larry now is ready to send encrypted e-mail. To do so, he composes an e-mail normally. When he is ready to send, he clicks Security on the menubar and selects PGP Encrypt. When he clicks Send, Larry is asked for his GPG passphrase. Once entered, the message is sent. Before I move on, I should mention that you can select PGP Sign from the Security menu if all you want to do is sign the message.

As wonderful as all this encryption and decryption stuff is, there are some problems. For instance, many packages (including KMail) will encrypt the message in-line, but a few others, such as Evolution, do not. Instead, encrypted messages are MIME attachments. Consequently, reading a message from KMail (or Eudora or Outlook) in Evolution requires that you save the message as a text file. You then can decrypt it with this:

```
gpg -d message.txt
```

To read an Evolution-encrypted message sent to KMail (and a few others) you need to save the attachment as opposed to the message itself. Decrypting it is the same process as above.

Many e-mail clients are available in the Linux world. KMail and Evolution are popular graphical choices, but so is Mozilla. For users of Mozilla, a plugin called Enigmail allows for seamless encryption and signing of messages. Text-based clients also support encryption, either on their own or through plugins; mutt and pine are popular examples.

Mon Dieu! Has the time gone by so quickly? I fear, *mes amis*, that closing time is upon us once again. François will refill your glasses once more before you go. As you leisurely sip that last glass, take some time to create and exchange public keys with one another. Perhaps I may be convinced to share the *Crème Linuxaise* with some of you—assuming proper security precautions, of course. Who knows, our security may be the very reason that door in the wine cellar remains tightly shut! Until next time, *mes amis*, let us all drink to one another's health. *A vôtre santé! Bon appétit!*

Marcel Gagné (mggagne@salmar.com) lives in Mississauga, Ontario. He is the author of *Moving to Linux: Kiss the Blue Screen of Death Goodbye!* (ISBN 0-321-15998-5) from Addison Wesley. His first book is the highly acclaimed *Linux System Administration: A User's Guide* (ISBN 0-201-71934-7). In real life, he is president of Salmar Consulting, Inc., a systems integration and network consulting firm.

RESOURCES

Gnu Privacy Guard (GnuPG): **www.gnupg.org**

KMail: **kmail.kde.org**

Mozilla: **www.mozilla.org**

Mozilla Enigmail: **enigmail.mozdev.org**

Ximian Evolution: **www.ximian.com/products/evolution**

Marcel's Web Site (check out the wine page): **www.marcelgagne.com**

www.asacomputers.com

Hardware Systems for the Open Source Community—since 1989
(Linux, FreeBSD, NetBSD, OpenBSD, Solaris...)

4TB in 4U—$7,599
Intel E7501, 2 x Intel Xeon 2.4GHz
1 GB ECC DDR Max: 16GB
1 x 10/100 LAN
1 Gigabit LAN
Raid 0,1,5,etc.
CD, Floppy
Redundant Hot-Swap Power Supply

1U Dual Itanium IDE—$4,889
Dual Intel Itanium2 1.4 Ghz
2 GB ECC DDR
1 of 4 x 40 GB HDD
Dual Gigabit LAN
Based on Supermicro 6113M-i

Dual AMD Opteron in 1U—$1,659
2 x AMD Opteron 240
1 GB DDR ECC Registered Memory
1 x 40 GB IDE
Gigabit LAN

AMD Opteron™

Space-saving Mini-1U Dual Xeon IDE—$1,249
1U 14" Depth
533 MHz FSB
1 of 2 x Intel Xeon 2.4 GHz
512 MB DDR ECC Memory
40 GB IDE HDD
1 GIG, 1 10/100 LAN

1U P4 Entry-level Server—$735
1 x Intel P4 2.4 GHz
256 MB DDR ECC Memory
40 GB IDE Drive
CD, 2 x 10/100 LAN

1U Dual Xeon SCSI—$1,577
SuperMicro 6013P-8+ Barebones
1 of 2 Intel Xeon 2.4 GHz 533 FSB
512 MB DDR ECC Max: 12GB
36 GB 10K rpm SCSI
4 Hot-Swap SCSI Drive Bays
CD, Floppy
Dual Gig, VGA onboard
Raid optional

ASA Colocation
$50 per month for
1U Rack - 20 GB/month

ASA Colocation Special
First month of colocation free.*

Storage Solutions
IDE, SCSI, Fiber RAID solutions
TB storage options
3Ware, Promise, Adaptec,
ICP-Vortex, LSI, CMD, Chaparral,
Mylex controllers
JMR, Kingston/Storcase solutions

Clusters
Rackmount and Desktop nodes
HP, Intel, 3Com, Cisco switches
KVM or Cyclades Terminal Server
APC or Generic racks

All systems installed and tested with user's choice of Linux distribution (free). ASA Colocation—$50 per month

ASA Computers, Inc.
1-800-REAL-PCS
2354 Calle Del Mundo, Santa Clara, CA 95054
TEL: (408) 654-2900 • FAX: (408) 654-2910
E-MAIL: sales@asacomputers.com
*Free delivery in Silicon Valley**
Open on Sat. too—10 to 1
Prices and availability of systems subject to change without notice
** Certain restrictions apply.*

www.asacomputers.com
www.b2bforsale.com

Application Proxying with Zorp, Part II

The Zorp proxy server works with the kernel Netfilter to make an application-level proxy that looks transparent to the client. **BY MICK BAUER**

In my last column, I sang the praises of application-layer proxy firewalls and introduced Balazs Scheidler's Zorp firewall suite, available in both commercial and free-of-charge versions. This column continues where we left off, discussing basic Zorp configuration for a simple inside-DMZ-outside scenario. We are going to configure only a couple of services, but this should be enough to help prospective Zorp users begin building their own intelligent firewall systems.

To review, application-layer proxies broker rather than merely pass the traffic that flows through them. For example, when a user on one network initiates an HTTP session on the other side of a proxying firewall, the firewall intercepts and breaks the connection, acting both as the server (from the client's viewpoint) and as the client (from the destination server's standpoint).

Zorp uses transparent proxies, which means that users behind a Zorp firewall need not be aware that the firewall is there; they may target foreign addresses and hostnames without configuring their software to communicate with the proxy. This is an important mitigator against the ugly fact that proxies are inherently more complicated than other kinds of firewalls. With Zorp, all the complexity is in the back end, resulting in much happier end users.

But that doesn't mean Zorp is painful for its administrators, either. I'd rate its complexity as being higher than iptables but lower than sendmail.cf. So without further ado, let's configure ourselves a Zorp firewall.

Assumptions

This article assumes that, per my last column, you've successfully patched your Linux 2.4 kernel and your iptables binary to support the TPROXY module (see www.balabit.com/products/oss/tproxy). It also assumes you have compiled and/or installed packages for libzorpll, zorp and zorp-modules; source code and deb packages are available at www.balabit.com/products/zorp_gpl. My examples further assume you're running Zorp GPL version 2.0, though the examples should apply equally to Zorp Pro 2.0. Zorp Pro has some proxy modules not included with Zorp GPL, but the modules common to both behave the same.

The Scenario

Zorp supports many more than three interfaces per firewall, but the most common firewall architecture nowadays is the three-homed-host architecture shown in Figure 1. This is the architecture I cover here.

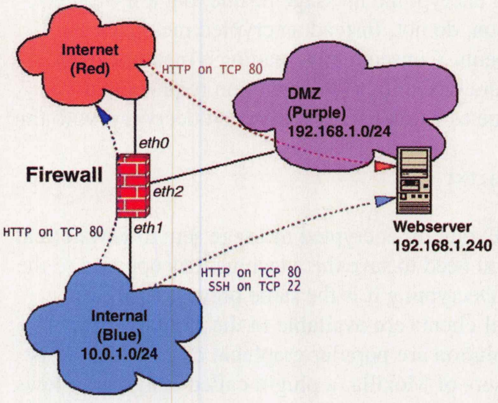


Figure 1. Example Architecture

Similarly, as you can see in Figure 1, we've got only three data flows: HTTP from the Internet to a DMZed Web server; HTTP from the internal network to the Internet; and HTTP and SSH from the internal network to the DMZ. Absent are things like IMAP, NNTP, FTP and other services that even simple setups commonly use. If you understand how to configure Zorp to accommodate these, though, you should be able to figure out others. I do, however, discuss DNS and SMTP, even though I omitted them from Figure 1.

Configuring a Dummy Interface

The first thing we need to do doesn't directly involve Zorp but rather the TPROXY kernel module. In transparent proxying, TPROXY needs a dummy network interface to bind to whenever it splits a data flow in two. This needs to be an interface whose IP address is neither Internet-routable nor associated with any network connected to the firewall.

Linux 2.4 kernels compile with support for dummy network interfaces by default. You should have one, unless you intentionally compiled your kernel without dummy driver support. If so, compile a new kernel with dummy support. All you need to do for TPROXY's purposes, therefore, is explicitly configure dummy0 with a nonroutable and unused address. In Debian, you should add the following lines to /etc/networking/interfaces:

```
auto dummy0
iface dummy0 inet static
     address 1.2.3.4
     netmask 255.255.255.255
```

Other distributions handle network configuration differently—Red Hat and SuSE use ifcfg- files in /etc/sysconfig/network—but hopefully you get the picture. Notice the 32-bit network mask: I repeat, this address must not belong to a real network.

iptables Configuration

You may be wondering, isn't this article about Zorp and not iptables? Yes, but Zorp runs in conjunction with iptables, not in place of it. TPROXY, in fact, is specifically a Netfilter patch.

LINUX CLUSTERS
THE HPC REVOLUTION

LCI LINUX CLUSTERS INSTITUTE

5th International Conference on Linux Clusters: The HPC Revolution 2004

May 18-20, 2004
May 17, 2004 Tutorials
Texas Advanced Computing Center
The University of Texas at Austin
Austin, Texas

http://www.linuxclustersinstitute.org/Linux-HPC-Revolution/

The 5th International Conference on Linux Clusters: The HPC Revolution 2004 is the premier international forum designed for:

- Fostering interactions among people who are working with Linux clusters.
- Building the Linux High-Performance cluster community.

Keynote Speakers

David Jursik
V. P. of IBM Worldwide Deep Computing Sales

Dr. Reza Rooholamini
Director of Enterprise Solutions, Dell Product Group

Dr. Edward Seidel
Director of LSU's Center for Applied Information Technology and Learning (CAPITAL).

This world-class event is focused on the application, management, and scientific computing techniques for high-performance Linux clusters.

Users, researchers, vendors, executives, and developers like you will gather in Austin, Texas, to learn the latest news and techniques while also helping others by sharing their own knowledge and experience.

The conference will feature speakers from academia, research labs, and industry involved in cluster-based high-performance computing.

Registration open: March 1, 2004

Early registration deadline: April 2, 2004

For detailed information about the conference program and registration, please see the conference web site at: http://www.linuxclustersinstitute.org/Linux-HPC-Revolution/

To use TPROXY, we need to configure it with the iptables command, as we do for the rest of Netfilter. (Netfilter is the proper name for Linux 2.4's firewall code—iptables is its front-end command.)

In addition, it's recommended that you run certain services, namely DNS and SMTP, on the firewall as self-contained proxies. If you do, you need to use iptables to configure your firewall to accept those connections directly. For example, BIND v9 supports split-horizon DNS, in which external clients are served from different zone files than are internal clients. Similarly, Postfix is easy to configure to act as a relay on behalf of internal hosts, but strictly as a local deliverer when dealing with external hosts. It makes sense to run such proxy-like services on a firewall, as long as you configure them extremely carefully.

If you're new to Netfilter/iptables, what follows may make little sense, and space doesn't permit me to explain it all in detail. Zorp is, after all, an advanced tool. In a nutshell, what we're going to do with iptables is run all packets through some simple checks against spoofed IP addresses. We then are going to intercept packets that need to be proxied transparently and process them in custom chains rather than by using the normal FORWARD chain. Technically, nothing is forwarded. Finally, we pass some packets that are destined for the firewall itself.

Zorp Pro includes a group of scripts collectively called iptables-utils, which simplify iptables management for Zorp. A free version of iptables-utils for Zorp GPL 2.0 is available at www.balabit.com/downloads/zorp/zorp-os/pool/i/iptables-utils. I highly recommend iptables-utils, as it makes it much easier to test a new iptables configuration before actually committing it.

Because it uses a syntax that I don't have space here to explain, the following example is instead a conventional iptables startup script. Here are the most important parts of such a script. First should come rules for the special tproxy table that the TPROXY module adds to Netfilter (Listing 1). This is where we define a custom proxy chain for each of our networks: PRblue for proxied connections initiated from our internal network; PRpurple for proxied connections initiated from our DMZ (none, in this scenario); and PRred for proxied connections originating from the Internet.

Several things are worth pointing out in Listing 1. First, notice that the tproxy table contains its own PREROUTING and OUTPUT output chains. We use the tproxy/PREROUTING chain in Zorp to route packets to the proper custom proxy chain (PRblue), based on the interface each packet enters. As with any custom iptables chain, if a packet passes through one of these without matching a rule, it's sent back to the line immediately following the rule that sent the packet to the custom chain. This is why custom chains don't have default targets.

In the PRblue chain, we've got two rules, one for each type of transaction allowed to originate from the internal network. All outbound HTTP material is proxied, that is, handed to a proxy process listening on port 50080. But in the SSH rule, we tell Netfilter to proxy all outbound SSH traffic unless it's destined for the firewall itself. Although Figure 1 doesn't show such a data flow (Blue→SSH→firewall), we need it in order to administer the firewall. This flow also requires a rule in the regular filter table's INPUT chain. In this example scenario, our DMZed Web server isn't permitted to initiate any connections itself, so we've created a PRpurple chain without actually populating it.

Listing 1. TPROXY Rules

```
iptables -t tproxy -P PREROUTING ACCEPT
iptables -t tproxy -A PREROUTING -i eth1 -j PRblue
iptables -t tproxy -A PREROUTING -i eth2 -j PRpurple
iptables -t tproxy -A PREROUTING -i eth0 -j PRred

iptables -t tproxy -P OUTPUT ACCEPT

iptables -t tproxy -N PRblue
iptables -t tproxy -A PRblue -p tcp --dport 80 \
  -j TPROXY --on-port 50080
iptables -t tproxy -A PRblue -p tcp --dport 22 \
  ! -d firewall.example.net -j TPROXY --on-port 50022

iptables -t tproxy -N PRpurple

iptables -t tproxy -N PRred
iptables -t tproxy -A PRred -p tcp --dport 80 \
  -j TPROXY --on-port 50080
```

Listing 2. Filter Table INPUT Chain

```
iptables -P INPUT DROP
iptables -A INPUT -j noise
iptables -A INPUT -j spoof
iptables -A INPUT -m tproxy -j ACCEPT
iptables -A INPUT -m state \
  --state ESTABLISHED,RELATED -j ACCEPT
iptables -A INPUT -i lo -j ACCEPT
iptables -A INPUT -i eth1 -j LOblue
iptables -A INPUT -i eth0 -j LOred
iptables -A INPUT -i eth2 -j LOpurple
iptables -A INPUT -j LOG --log-prefix "INPUT DROP: "
iptables -A INPUT -j DROP
```

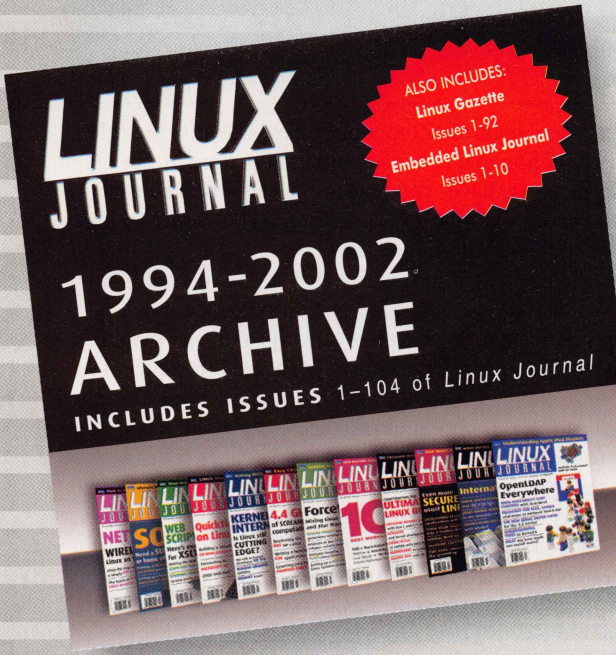
Now we move on to the regular filter table, this is the Netfilter table most of us are used to dealing with—it's the default when you omit the -t option with iptables. Listing 2 shows our example firewall's filter table's INPUT rules.

The first few lines check packets against some custom chains that check for spoofed IP addresses; if they pass those checks, they continue down the INPUT chain. Packets generated by the TPROXY module itself are accepted, as are packets belonging to established allowed transactions and loopback packets (lines 4–6, respectively). Next, as with the tproxy table's PREROUTING chain, we route packets to custom chains based on ingress interface. This time, the custom chains are for packets with local destinations, as opposed to proxied ones, so I've named them LOblue and so forth. Next come our filter table's custom chains (Listing 3).

The first three of these custom chains are the most important: LOblue, LOpurple and LOred tell Netfilter how to process packets destined for the firewall itself, based on in which interface the packets arrive. In LOblue, we're accepting DNS queries, SSH connections and SMTP connections. In LOpurple, we're accepting only DNS queries. And in LOred, we're accepting DNS replies from our ISP's DNS server (upstream.dns.server)

Missing an issue?

In easy-to-use HTML format, the space-saving 1994-2002 *Linux Journal* Archive CD-ROM offers users the advantage of immediate access to an essential Linux resource: *Linux Journal*.


LINUX JOURNAL
1994-2002
ARCHIVE
INCLUDES ISSUES 1–104 of Linux Journal
ALSO INCLUDES:
Linux Gazette
Issues 1-92
Embedded Linux Journal
Issues 1-10

Inside each issue of *Linux Journal* are the great features we've become so well-known for:

- **Time-saving programming tricks & tips**
- **Avoid common mistakes by following *LJ*'s thorough how-tos**
- **Interviews with the industry's leading personalities**
- **Concise product reviews**

The Archive CD-ROM contains every issue of *Linux Journal*, from the premiere March 1994 issue through December 2002. Issues 1-92 of *Linux Gazette* and 1-10 of *Embedded Linux Journal* are also included.

Over 100 issues for only $29.95!

Available only at the Linux Journal Store:

http://store.linuxjournal.com/

Listing 3. Custom Chains in the Filter Table

```
iptables -N LOblue
iptables -A LOblue -p tcp --dport 22 --syn -j ACCEPT
iptables -A LOblue -p udp --dport 53 -j ACCEPT
iptables -A LOblue -p tcp --dport 25 --syn -j ACCEPT
iptables -A LOblue -j LOG --log-prefix "LOblue DROP: "
iptables -A LOblue -j DROP

iptables -N LOpurple
iptables -A LOpurple -p udp --dport 53 -j ACCEPT
iptables -A LOpurple -j LOG \
  --log-prefix "LOpurple DROP: "
iptables -A LOpurple -j DROP

iptables -N LOred
iptables -A LOred -p udp -s upstream.dns.server \
  -sport 53 -j ACCEPT
iptables -A LOred -p tcp --dport 25 --syn -j ACCEPT
iptables -A LOred -j LOG --log-prefix "LOred DROP: "
iptables -A LOred -j DROP

iptables -N noise
iptables -A noise -p udp --dport 137:139 -j DROP
iptables -A noise -j RETURN

iptables -N spoof
iptables -A spoof -i lo -j RETURN
iptables -A spoof ! -i lo -s 127.0.0.0/8 -j spoofdrop
iptables -A spoof -i eth1 ! -s 10.0.1.0/24 \
  -j spoofdrop
iptables -A spoof ! -i eth1 -s 10.0.1.0/24 \
  -j spoofdrop
iptables -A spoof -i eth2 ! -s 192.168.1.0/24 \
  -j spoofdrop
iptables -A spoof ! -i eth2 -s 192.168.1.0/24 \
  -j spoofdrop
iptables -A spoof -j RETURN

iptables -N spoofdrop
iptables -A spoofdrop -j LOG \
  --log-prefix "Spoofed packet: "
iptables -A spoofdrop -j DROP
```

and SMTP connections. The last three of these custom chains are the simplest: noise filters NETBIOS packets, those notorious clutterers of Linux firewall logs; spoof filters for packets with obviously spoofed, that is, impossible, source IP addresses; and spoofdrop logs and drops packets caught by the spoof chain.

Listing 4 shows the remainder of our example iptables script, an essentially empty FORWARD chain with a default DROP policy and an empty OUTPUT chain with a default ACCEPT chain. Again, this is a proxying firewall, so it won't forward anything. You may be uneasy with the default ACCEPT policy for firewall-originated packets, but this is both necessary and safe on a Zorp firewall.

Listing 4. The Filter Table's FORWARD and OUTPUT Chains

```
iptables -P FORWARD DROP
iptables -A FORWARD -j LOG \
  --log-prefix "FORWARD DROP: "
iptables -A FORWARD -j DROP

iptables -P OUTPUT ACCEPT
```

Listing 5. instances.conf

```
blue -v3 -p /etc/zorp/policy.py \
  --autobind-ip 1.2.3.4
purple -v3 -p /etc/zorp/policy.py \
  --autobind-ip 1.2.3.4
red -v3 -p /etc/zorp/policy.py \
  --autobind-ip 1.2.3.4
```

Configuring Zorp's Instances

Finally, we come to actual Zorp configuration files. These are stored in /etc/zorp, and the first one we tackle is instances.conf, which defines and controls Zorp's instances. Usually, the rule of thumb is to define one instance per network zone, so in our example scenario we have, you guessed it, one instance each for our red, purple and blue zones. Listing 5 shows what such an instances.conf file would look like.

The first field in each line is the name of the instance. This is user-definable, but we need to refer to it verbatim in the Zorp configuration file proper, policy.py. Speaking of which, you may use separate configuration files for each instance if you wish, or you may configure multiple zones within a single file. Regardless, the -p option in instances.conf tells Zorp which file to use for each instance.

The -v parameter sets log message verbosity: 3 is the medium setting, and 5 is useful for debugging. This parameter controls only Zorp-generated log messages and has no effect whatsoever on Netfilter/iptables logging. Finally, each line ends with an --autobind-ip setting that determines to which dummy IP Zorp should bind TPROXY when proxying connections. This IP address can and should be shared between all instances. This address, obviously, should be the one you set earlier (see Configuring a Dummy Interface, above).

Configuring Zorp's Application Proxies: policy.py

Your iptables script determines how packets get routed to proxies, and /etc/zorp/instances.conf determines how Zorp starts up. But to tell Zorp's proxies how to behave, you need to set up /etc/zorp/policy.py, or whatever you called the configuration file(s) referenced in instances.conf—policy.py is conventional but not mandatory. This policy file contains two parts. The first part is a global section in which zones are defined based on network addresses and allowed services. The second part is a service-instance definition section in which each instance listed in instances.conf is defined based on which services originate in each and in which those services are mapped to application proxies.

Listing 6 shows a complete global section from our example policy.py. It begins with some import sections, in which essential Python functions are included. Next come our zone definitions. If you set up instances.conf to run one Zorp instance per zone, your zone names here can be similar to or even the same as your instance names. In Listing 6 I've chosen

BEST PRICE. BEST PERFORMANCE. BEST SUPPORT.

ImageStream
TransPort
Power
ACT/LINK
10/100
E0 E1 E2

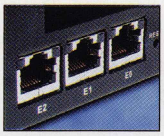
Evolution OF THE Revolution

The new TransPort™ has quickly become ImageStream's best-selling router. With its small footprint, business-class features, and competitive price, the TransPort is an ideal router for T1 and E1 applications that demand wire-speed performance.

The TransPort includes 128 MB RAM, three 10/100 ethernet ports, and an expansion slot for add-on cards. It also features ImageStream's Enterprise Linux™ router software that supports most WAN protocols, and advanced features including NAT firewall, packet filtering, bridging, bandwidth limiting, QoS, dynamic routing, VPN, and more.

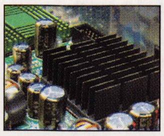
NO HASSLE 31 DAY RETURN GUARANTEE

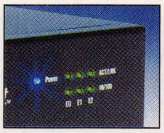
Like all ImageStream routers, the TransPort includes 12 months of free 24/7 technical support, a full 12-month warranty on parts and labor, free software upgrades for life, and the industry's only money back performance guarantee.



Well Connected The TransPort provides three 10/100 ethernet ports for flexible LAN connectivity.

Fanless CPU The TransPort's highly efficient CPU minimizes power consumption and cooling requirements.





Easy Indicators The TransPort provides front panel LEDs to simplify the analysis of LAN connections.

- **Best Software** Our Enterprise Linux™ router distribution has everything you need to deploy advanced network applications including secure shell, menu-based configuration, real-time monitoring, IP firewall, bridging, interface bonding, dynamic routing, QoS, IPsec VPN, free software upgrades for life, and more.
- **Best Support** ImageStream routers include a full year of free 24/7 support. When you consider total cost of ownership, ImageStream routers cost less up front and over the long haul.
- **Guaranteed Performance** Our 31-day performance guarantee ensures your router will function as promised. If your ImageStream router does not work as specified in writing, and our support team cannot correct the problem, you can return your router for a full refund of the purchase price. See our web site for details.

COMPARE.

Base Chassis
Three 10/100 ports • Firewall, filtering & VPN **$899**

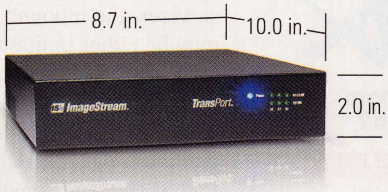
1-Port ADSL
One ADSL port • Three 10/100 ports • Firewall, filtering & VPN • Wire-speed performance* **$1,049**

1-Port T1/E1 or Serial
One T1/E1 or sync serial port • Three 10/100 ports • Firewall, filtering & VPN • Wire-speed performance* **$1,499**

2-Port T1/E1 or Serial
Two T1/E1 or sync serial ports • Three 10/100 ports • Firewall, bonding, filtering & VPN • Supports ATM & inverse multiplexing (IMA) • Wire-speed performance* **$1,999**

ImageStream
Rebel Router

ImageStream also offers the Rebel, which delivers line rate DS3/E3 performance at the industry's lowest price point.



Sized Right The TransPort's small footprint allows it to be installed just about anywhere.

ImageStream™
Routers for the Real World™

800.813.5123
www.imagestream.com
1.574.935.8484

**Note: ImageStream's wire-speed performance guarantee applies to WAN ports. Wire-speed specifications and notes are posted on the ImageStream web site.*

TransPort and "Routers for the Real World" are trademarks of ImageStream Internet Solutions, Inc. Linux is a registered trademark of Linus Torvalds. Specifications are subject to change without prior notice.
©2004 ImageStream Internet Solutions, Inc.

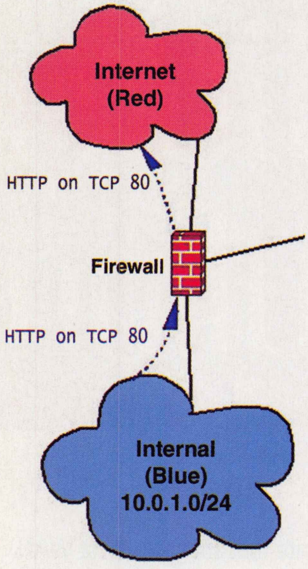


Figure 2. An HTTP Transaction Outbound from Blue, Inbound to Red

different names in order to illustrate that technically, zone names are distinct from instance names.

In each zone definition, you can see a network address that corresponds to those in Figure 1 and specifications of which services are allowed. These service names are user-definable and fleshed out in the subsequent service-instance definitions. The important thing to understand about these statements is that inbound and outbound is relative to the zone/network, not to the firewall.

Figure 2 shows what the internal-to-Internet HTTP data flow looks like as a proxied connection. In this illustration, we see this data flow exists both as an outbound connection out of the Internal (blue) zone and an inbound connection to the Internet (red) zone. This is borne out in the respective bluezone and redzone definitions in Listing 6. It's also important to use the same service name in both zone definitions that a given data flow traverses (blue_http in the case of Figure 2 and Listing 6).

Listing 6. policy.py, Part I (Global Settings)

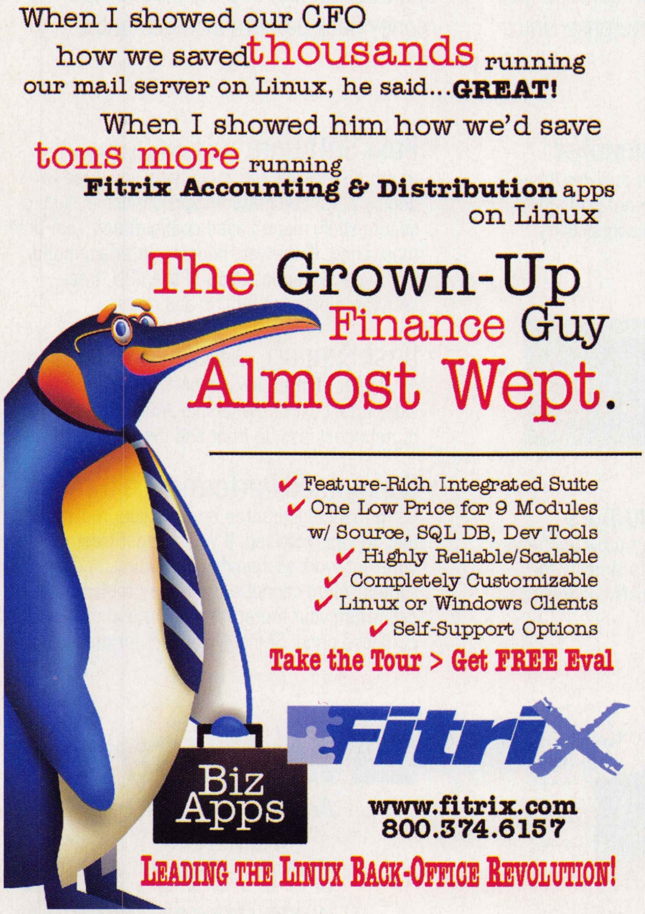
```
from Zorp.Core import *
from Zorp.Plug import *
from Zorp.Http import *

InetZone("bluezone", "10.0.1.0/24",
   outbound_services=["blue_http", "blue_ssh"],

InetZone("purplezone", "192.168.1.0/24",
   inbound_services=["blue_http", "blue_ssh",
                    "red_http"])

InetZone("redzone", "0.0.0.0/0",
   outbound_services=["red_http"],
   inbound_services=["*"])

InetZone("localzone", "127.0.0.0/8",
   inbound_services=["*"])

# end global section
```

Listing 7. policy.py, Part II (Instance Definitions)

```
def blue():
    Service("blue_http", HttpProxy,
            router=TransparentRouter())
    Service("blue_ssh", PlugProxy,
            router=TransparentRouter())
    Listener(SockAddrInet(´10.0.1.254´, 50080),
             "blue_http")
    Listener(SockAddrInet(´10.0.1.254´, 50022),
             "blue_ssh")

def purple():
    pass

def red():
    Service("red_http", HttpProxy,
router=DirectedRouter(SockAddrInet(´192.168.1.242´, 80),
forge_addr=TRUE))
    Listener(SockAddrInet(´169.254.1.254´, 50080),
             "red_http")
```

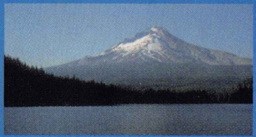
The last point to make about Listing 6 is the * wild card signifies all defined services. This is narrower than it might seem; * includes only those services defined in policy.py's service-instance definitions, not all possible services. Remember, Zorp processes only those packets that Netfilter and TPROXY send to it. If a given zone is to allow no outbound or inbound services, the inbound_services or outbound_services parameter may be either omitted or set to [] (empty brackets).

Listing 7 shows our policy.py file's service-instance definitions. The first line of each definition must reference an instance name specified in instances.conf, and the following lines in the definition must be indented because these rules are processed by Python, which is precise about indentation. The definition can't


When I showed our CFO how we saved thousands running our mail server on Linux, he said...GREAT!
When I showed him how we'd save tons more running Fitrix Accounting & Distribution apps on Linux
The Grown-Up Finance Guy Almost Wept.
✔ Feature-Rich Integrated Suite
✔ One Low Price for 9 Modules w/ Source, SQL DB, Dev Tools
✔ Highly Reliable/Scalable
✔ Completely Customizable
✔ Linux or Windows Clients
✔ Self-Support Options
Take the Tour > Get FREE Eval
Fitrix
Biz Apps
www.fitrix.com
800.374.6157
LEADING THE LINUX BACK-OFFICE REVOLUTION!

The Compilers and Tools for Scientists and Engineers.

C, C++, F77, F90 and HPF • 32-bit and 64-bit Linux
Optimized for Intel and AMD CPUs • 64-bit support on AMD64
Workstation, Server and Cluster configurations • Fast compile times
Native OpenMP • Native SMP auto-parallelization
Native distributed-memory parallelization • Cache tiling
Function inlining • SSE/SSE2 Vectorization • Loop unrolling
Interprocedural optimization • Profile-feedback optimization
Large file support on 32-bit Linux • 64-bit integers and pointers
F77 pointers • Byte-swapping I/O • VAX and IBM extensions
OpenMP/MPI/threads debugging • OpenMP/MPI/threads profiling
Interoperable with g77/gcc/gdb • PDF and printed documentation
Electronic purchase, download and upgrades • Tech support
Network-floating licenses • Academic and volume discounts



The Portland Group™ Compiler Technology

www.pgroup.com ++01 (503) 682-2806 sales@pgroup.com

The registered trademarks and marks are the property of their respective owners.

STMicroelectronics

be empty: if no services originate in a given instance, the token pass may be used, as with the purple() instance definition in Listing 7.

Otherwise, the definition should consist of one or more Service lines, specifying a service name referenced in one or more zone definitions, and a Zorp proxy module, either a built-in proxy included in the global import statements or defined in a custom class. The last field in a Service line is a router, which specifies where proxied packets should be sent. You can see in Listing 7 that for the red_http service, we've used the forge_addr=TRUE option to pass the source IPs of Web clients intact from the Internet to our Web server. Without this option, all Web traffic hitting the DMZ appears to originate from the firewall itself.

Although in Listing 7 we're using only the HttpProxy and the PlugProxy (a general-service UDP and TCP proxy that copies application data verbatim), Zorp GPL also has proxies for FTP, whois, SSL, telnet and finger. As I mentioned before, you also can create custom classes to alter or augment these proxies. It's easy to create, for example, an HTTP proxy that performs URL filtering or an SSL proxy stacked on an HTTP proxy so HTTPS traffic can be proxied intelligently. Unfortunately, these are advanced topics I can't cover here; fortunately, all of Zorp's Python proxy modules are heavily commented.

The TransparentRouter referenced in Listing 7 simply proxies the packets to the destination IP and port specified by the client. But in the red instance's red_http service, we see that a DirectedRouter, which requires a mandatory destination IP and port, may be specified instead.

Each Service line in a service-instance definition must have a corresponding Listener line. This line tells Zorp to which local (firewall) IP address and port the service should be bound. It may seem counterintuitive that the ports specified in Listing 7's Listener statements are high ports: 50080 instead of 80 and 50022 instead of 22. But remember, each proxy receives its packets from the kernel through Netfilter, not directly from clients. Accordingly, these high ports must correspond to those specified in your tproxy table Netfilter rules (Listing 1).

I mentioned that unlike HttpProxy, which is a fully application-aware proxy that enforces all relevant Internet RFCs for proper HTTP behavior, PlugProxy is a general-service proxy (GSP). Using PlugProxy still gives better protection than does packet filtering on its own, because the very act of proxying, even without application intelligence, insulates your systems from low-level attacks that Netfilter may not catch on its own.

Conclusion

And with that, we've scratched the dense surface of Zorp GPL. This is by far the most complex tool I've covered in these pages, but I think you'll find Zorp to be well worth the time you invest in learning how to use it.

Mick Bauer, CISSP, is *Linux Journal*'s security editor and an IS security consultant in Minneapolis, Minnesota. He's the author of *Building Secure Servers With Linux* (O'Reilly & Associates, 2002).

Looking for a GUI for your next embedded project?

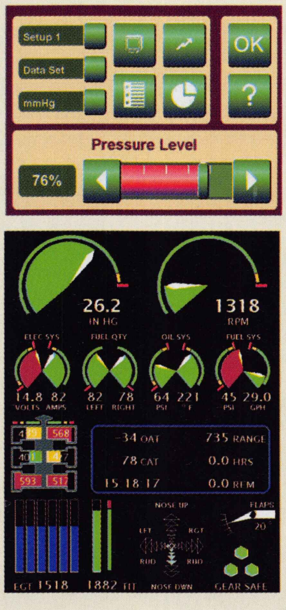
Look no further...

- Skilled and timely support to users around the globe.
- PEG includes a full set of Development Tools including WindowBuilder, ImageConvert, and FontCapture.
- PEG is fully customizable making your User Interface completely unique and distinguishing your product from your competitors.
- PEG is fully integrated with a wide selection of RTOSs and supports all popular target processors, video controllers and I/O devices.
- Small Footprint, Fast Execution Speed, Completely Rommable.
- PEG is delivered with Full Source Code.
- Multi-lingual support – including 2-byte character sets and UNICODE string encoding.
- Event-driven programming model – familiar to Windows developers.
- Free Evaluation Packages available for Desktop Prototyping.
- Unlimited seats and Royalty FREE!!

Use only the best GUI Library and Development Tools on your next project.

PEG is used by Fortune 500 Companies around the globe. Join the growing list of Software Engineers that are using PEG to create some of the most exciting products on the market!

PEG™



SWELL SOFTWARE Inc.

Free User's Manual or to obtain a FREE PEG Evaluation Kit, **call (810) 982-5955 or visit www.swellsoftware.com.**

RESOURCES

The English-language home for Balabit, creators of Zorp: **www.balabit.com**.

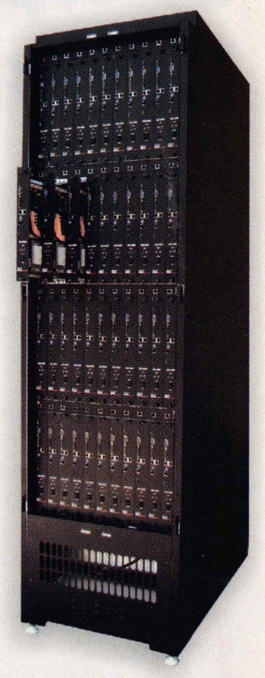
The root download directory for ZorpOS contains some tools that make using Zorp GPL much easier, including iptables-utils, a TPROXY-enabled Linux kernel and iptables command. In fact, these are the free parts of the Debian distribution included with Zorp Pro, which is why everything in ZorpOS is in the form of Debian packages. If you aren't a Debian user, everything you want is in the subdirectories of pool; at the top of each package's subdirectory are tar.gz files containing source code. If you are a Debian user, you can use the URL as an apt-get source: **www.balabit.com/downloads/zorp/zorp-os**.

The Zorp Users' Mailing List is an amazingly quick and easy way to get help using Zorp, whether Pro or GPL. This URL is the site for subscribing to it or browsing its archives. Note that Balabit is a Hungarian company and its engineers (and some of the most helpful Zorp users) operate in the CET (GMT+1) time zone: **https://lists.balabit.hu/mailman/listinfo/zorp**.

Experience the most computing power in the smallest footprint possible...

The RackSaver® BladeRack™, powered by Intel® Xeon® processors, is designed to support a wide variety of enterprise solutions...including yours.

Discover why industry leaders like Industrial Light and Magic, Pixar Animation Studios, NVIDIA, Shell, WesternGeco, Veritas DGC, and Novartis rely on RackSaver for their high performance computing needs. RackSaver has provided customers in every industry from Entertainment and Electronic Design Automation to Oil and Gas with the most powerful, reliable, and scalable high performance clusters available on the market today. Our customers understand that by increasing density in their server room they improve productivity while lowering their overall total cost of ownership.

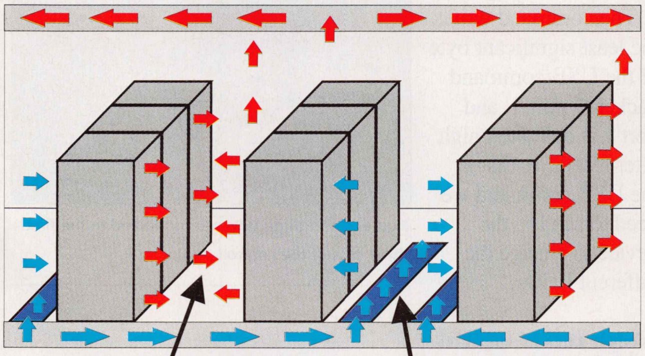


The RackSaver BladeRack Can House up to 176 Intel Xeon Processors in a Standard Rack Footprint

Discover the RackSaver Advantage - Our Patented Vertical Cooling Technology™

RackSaver's Vertical Cooling Technology is what separates us from the competition. No one else offers it. A traditional cluster installation with front to back or front to top cooling can cause serious hot spots in data centers. Utilizing the RackSaver Patented Vertical Cooling System eliminates data center hot spots allowing for increased density in our BladeRack without overheating.

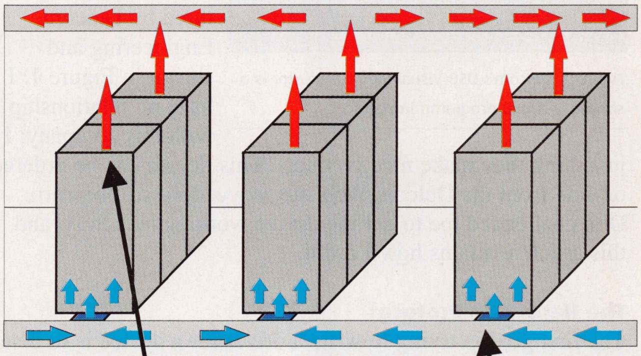
A Traditional Cluster Installation Can Cause Serious Hot Spots



Heat Exhausts From Racks and Hot Spots Occur

Cool Air Enters Room Near Racks

RackSaver's Patented Vertical Cooling Technology Eliminates Data Center Hot Spots



Heat Exhausts From the Hood of Each BladeRack Making Hot Spots Non-Existent

Cool Air Enters Intakes Located Under Each BladeRack Maximizing the Use of the Cold Air

RackSaver®
Leader in High-Density Servers!

Start configuring your company's high performance supercomputing solution now by visiting RackSaver online at www.racksaver.com or call one of our Account Managers at (888) 942-3800.

© 2004, RackSaver Inc. All Rights Reserved. RackSaver, the RackSaver logo, BladeRack, and Vertical Cooling Technology are trademarks or registered trademarks of RackSaver Inc. Intel and Xeon and combinations thereof are trademarks of Intel Corporation or its subsidiaries in the United States or other countries.

Writing a Simple USB Driver

Give your Linux box a multicolored light you can see from across the room, and learn how to write a simple driver for the next piece of hardware you want to hook up. **BY GREG KROAH-HARTMAN**

Since this column began, it has discussed how a Linux driver writer can create various types of kernel drivers, by explaining the different kernel driver interfaces including TTY, serial, I2C and the driver core. It is time to move on now and focus on writing real drivers for real hardware. We start by explaining how to determine what kind of kernel driver interface to use, tricks to help figure out how the hardware actually works and a lot of other real-world knowledge.

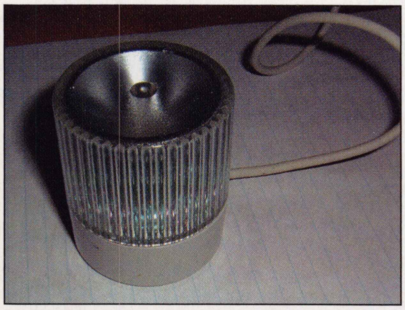


Figure 1. Delcom's USB Visual Signal Indicator is a simple first USB programming project.

Let's begin with a goal of making a simple USB lamp device work well with Linux. Editor Don Marti pointed out a neat device, the USB Visual Signal Indicator, manufactured by Delcom Engineering and shown in Figure 1. I have no relationship with this company; I just think they make nice products. This device can be ordered on-line from the Delcom Web site, www.delcom-eng.com. Don challenged me to get the device working on Linux, and this article explains how I did it.

The Hardware Protocol

The first goal in trying to write a driver for a device is to determine how to control the device. Delcom Engineering is nice enough to ship the entire USB protocol specification their devices use with the product, and it also is available on-line for free. This documentation shows what commands the USB controller chip accepts and how to use them. They also provide a Microsoft Windows DLL to help users of other operating systems write code to control the device.

The documentation for this device is only the documentation for the USB controller in the lamp. It does not explicitly say how to turn on the different color LEDs. For this, we have to do a bit of research.

NO DOCS? REVERSE ENGINEER IT!

If the USB protocol for this device had not been documented or available to me, I would have had to reverse engineer this information from the device itself. A handy tool for this kind of work is a free program called USB Snoopy, www.wingmanteam.com/usbsnoopy; another version of it is SnoopyPro, usbsnoop.sourceforge.net. These programs are both Windows programs that allow users to capture the USB data that is sent to and received from any USB device on a Windows system. All a developer needs to do is find a Windows machine, install the Windows driver provided by the manufacturer for the device and run the snoop program. The data is captured to a file to be analyzed later. Perl scripts can help filter some of the extra noise in the output of these snoop programs into an easier format to understand.

Another method a few people have used to reverse engineer the USB protocol of a device is to run a Windows instance using VMware on top of Linux. VMware enables the Windows instance to talk to all of the USB devices plugged in to the Linux machine by sending data to Linux though the usbfs. A simple modification to the usbfs causes all data flowing though it to be logged to the kernel log. Using this, the full USB traffic stream can be captured and later analyzed.

After opening up the lamp device, making sure not to lose the spring that easily pops out when unscrewing the device, the circuit board can be inspected (Figure 2). Using an ohmmeter, or any kind of device for detecting a closed circuit, it was determined that the three different LEDs are connected to the first three pins of port 1 on the main controller chip.

In reading the documentation, the USB command to control the levels of the port 1 pins is `Major 10, Minor 2, Length 0`. The command writes the least significant byte of the USB command packet to port 1, and port 1 is defaulted high after reset. So, that is the USB command we need to send to the device to change the different LEDs.

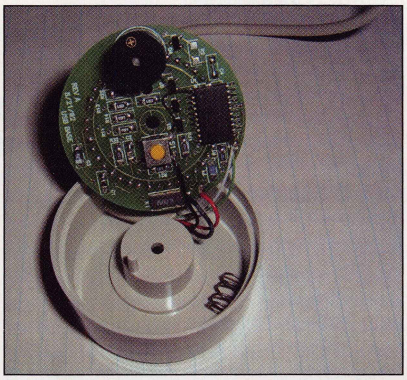
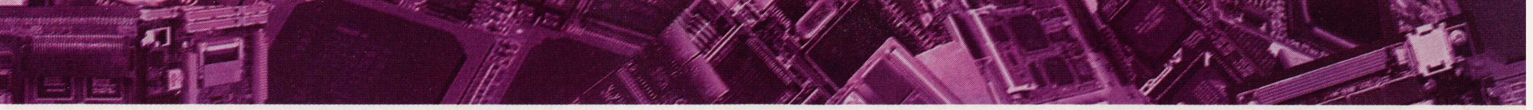


Figure 2. The three LEDs are connected to the first three pins of the controller chip.

Which LED Is Which?

Now that we know the command to enable a port pin, we need to determine which LED color is connected to which pin. This is easy to do with a simple program that runs through all possible combinations of different values for the three port pins and then sends the value to the device. This program enabled me to create a table of values and LED colors (Table 1).

So, if all pins on the port are enabled (a value of 0x07 hex), no LEDs are on. This matches up with the note in the data sheet that stated, "Port 1 is defaulted high after reset." It would make sense not to have any LEDs enabled when the device is first plugged in. This means we need to turn port pins low (off)



in order to turn on the LED for that pin. Using the table, we can determine that the blue LED is controlled by pin 2, the red LED by pin 1 and the green LED by pin 0.

A Kernel Driver

Armed with our new-found information, we set off to whip up a quick kernel driver. It should be a USB driver, but what kind of interface to user space should we use? A block device does not make sense, as this device does not need to store filesystem data, but a character device would work. If we use a character device driver, however, a major and minor number needs to be reserved for it. And how many minor numbers would we need for this driver? What if someone wanted to plug 100 different USB lamp devices in to this system? To anticipate this, we would need to reserve at least 100 minor numbers, which would be a total waste if all anyone ever used was one device at a time. If we make a character driver, we also would need to invent some way to tell the driver to turn on and off the different colors individually. Traditionally, that could be done using different ioctl commands on the character driver, but we know much better than ever to create a new ioctl command in the kernel.

As all USB devices show up in their own directory in the sysfs tree, so why not use sysfs and create three files in the USB device directory, blue, red and green? This would allow any user-space program, be it a C program or a shell script, to change the colors on our LED device. This also would keep us from having to write a character driver and beg for a chunk of minor numbers for our device.

To start out our USB driver, we need to provide the USB subsystem with five things:

- A pointer to the module owner of this driver: this allows the USB core to control the module reference count of the driver properly.
- The name of the USB driver.
- A list of the USB IDs this driver should provide: this table is used by the USB core to determine which driver should be matched up to which device; the hot-plug user-space scripts use it to load that driver automatically when a device is plugged in to the system.
- A probe() function called by the USB core when a device is found that matches the USB ID table.
- A disconnect() function called when the device is removed from the system.

Table 1. Port Values and the Resulting LED Patterns

Port value in hex	Port value in binary	LEDs on
0x00	000	Red, Green, Blue
0x01	001	Red, Blue
0x02	010	Green, Blue
0x03	011	Blue
0x04	100	Red, Green
0x05	101	Red
0x06	110	Green
0x07	111	No LEDs on

The driver retrieves this information with the following bit of code:

```
static struct usb_driver led_driver = {
        .owner = THIS_MODULE,
        .name =  "usbled",
```

Wireless Connectivity
- Cellular Modem
- 802.11b WiFi

TS-5600 586 Single Board Computer

Linux 2.4

$279 qty 1
Shown with optional IDE Compact Flash, A/D, DAC, RS-485 and Orinoco WiFi adapter.

- 133 MHz x86
- PCMCIA slot
- 10/100 Ethernet
- PC/104 Bus
- Compact Flash
- 32-64 MB SDRAM
- 6 ch 12-bit A/D
- 2 ch 12-bit DAC
- 23 Digital I/O
- 2 Serial Ports

- PC compatible: BIOS, DOS, RTC, I/O & memory map
- User programmable PLD for digital I/O
- Low power +5v @ 900mA
- 4.3" x 5.6"
- TS-Linux 3.0 is open source GNU/Linux - Glibc 2.2.5
- Apache, SSH, Perl, bash, FTP, kernel drivers

Technologic SYSTEMS

(480)-837-5200
www.embedded586.com

```
        .probe =        led_probe,
        .disconnect =   led_disconnect,
        .id_table =     id_table,
};
```

The id_table variable is defined as:

```
static struct usb_device_id id_table [] = {
        { USB_DEVICE(VENDOR_ID, PRODUCT_ID) },
        { },
};
MODULE_DEVICE_TABLE (usb, id_table);
```

The led_probe() and led_disconnect() functions are described later.

When the driver module is loaded, this led_driver structure must be registered with the USB core. This is accomplished with a single call to the usb_register() function:

```
retval = usb_register(&led_driver);
if (retval)
        err("usb_register failed. "
            "Error number %d", retval);
```

Likewise, when the driver is unloaded from the system, it must unregister itself from the USB core:

```
usb_deregister(&led_driver);
```

The led_probe() function is called when the USB core has found our USB lamp device. All it needs to do is initialize the device and create the three sysfs files, in the proper location. This is done with the following code:

```
/* Initialize our local device structure */
dev = kmalloc(sizeof(struct usb_led), GFP_KERNEL);
memset (dev, 0x00, sizeof (*dev));

dev->udev = usb_get_dev(udev);
usb_set_intfdata (interface, dev);

/* Create our three sysfs files in the USB
* device directory */
device_create_file(&interface->dev, &dev_attr_blue);
device_create_file(&interface->dev, &dev_attr_red);
device_create_file(&interface->dev, &dev_attr_green);

dev_info(&interface->dev,
    "USB LED device now attached\n");
return 0;
```

The led_disconnect() function is equally as simple, as we need only to free our allocated memory and remove the sysfs files:

```
dev = usb_get_intfdata (interface);
usb_set_intfdata (interface, NULL);

device_remove_file(&interface->dev, &dev_attr_blue);
device_remove_file(&interface->dev, &dev_attr_red);
device_remove_file(&interface->dev, &dev_attr_green);

usb_put_dev(dev->udev);
kfree(dev);

dev_info(&interface->dev,
        "USB LED now disconnected\n");
```

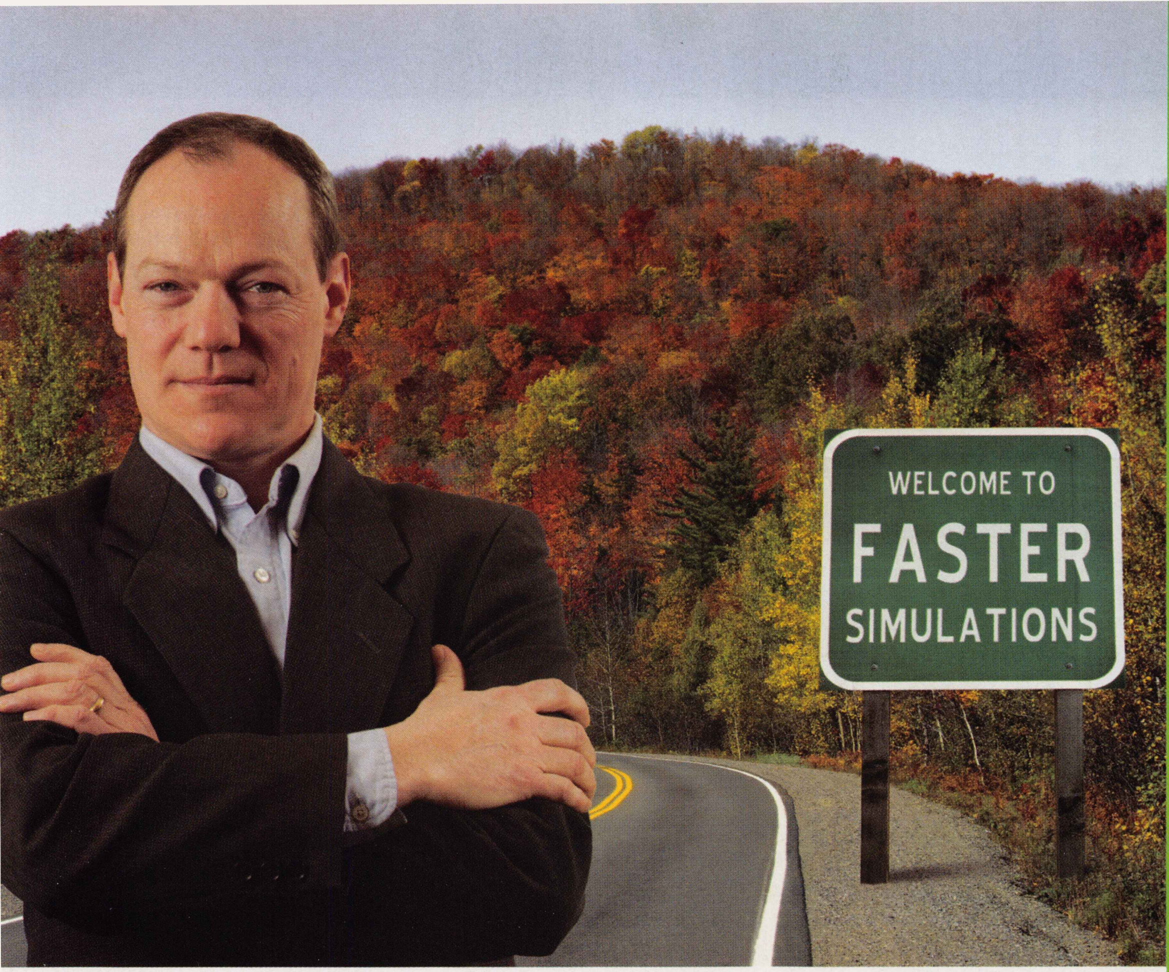
When the sysfs files are read from, we want to show the current value of that LED; when it is written to, we want to set that specific LED. To do this, the following macro creates two functions for each color LED and declares a sysfs device attribute file:

```
#define show_set(value)                                 \
static ssize_t                                          \
show_##value(struct device *dev, char *buf)             \
{                                                       \
   struct usb_interface *intf =                         \
      to_usb_interface(dev);                            \
   struct usb_led *led = usb_get_intfdata(intf);        \
                                                        \
   return sprintf(buf, "%d\n", led->value);             \
}                                                       \
                                                        \
static ssize_t                                          \
set_##value(struct device *dev, const char *buf,        \
             size_t count)                              \
{                                                       \
   struct usb_interface *intf =                         \
      to_usb_interface(dev);                            \
   struct usb_led *led = usb_get_intfdata(intf);        \
   int temp = simple_strtoul(buf, NULL, 10);            \
                                                        \
   led->value = temp;                                   \
   change_color(led);                                   \
   return count;                                        \
}                                                       \

static DEVICE_ATTR(value, S_IWUGO | S_IRUGO,
                       show_##value, set_##value);
show_set(blue);
show_set(red);
show_set(green);
```

This creates six functions, show_blue(), set_blue(), show_red(), set_red(), show_green() and set_green(); and three attribute structures, dev_attr_blue, dev_attr_red and dev_attr_green. Due to the simple nature of the sysfs file callbacks and the fact that we need to do the same thing for every different value (blue, red and green), a macro was used to reduce typing. This is a common occurrence for sysfs file functions; an example of this in the kernel source tree is the I2C chip drivers in drivers/i2c/chips.

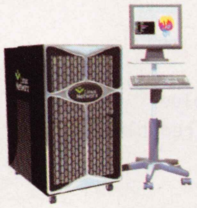
So, to enable the red LED, a user writes a 1 to the red file in sysfs, which calls the set_red() function in the driver, which calls the change_color() function. The change_color() function looks like:

```
#define BLUE            0x04
#define RED             0x02
#define GREEN           0x01
```



“From the design of the rack to the cluster management software, the level of refinement in the Linux Networx system was remarkable. We were very impressed and can recommend Linux Networx without any reservations.”

Paul Bemis
Vice President,
Fluent Inc.



You know where you want to go, Linux Networx can take you there.

Knowing where you want to go is easy. Reaching your destination quickly and smoothly can be more difficult. At Linux Networx we design comprehensive Linux cluster systems tailored to your specific needs so you can reach your destination in record time–without long delays or unexpected detours.

More than just hardware our Evolocity® cluster systems are optimized to run applications, like Fluent's CFD software, with the necessary horsepower and other essential software to achieve fast and accurate simulations. In fact, Fluent was so impressed, they use an Evolocity cluster system to power their Remote Simulation Facility (RSF).

case study

To read about the Linux Networx Fluent RSF cluster system and our Five Points of Proven Value, visit www.linuxnetworx.com/proven or call 1-877-505-5694.

FLUENT

Linux Networx™
proven

```
buffer = kmalloc(8, GFP_KERNEL);

color = 0x07;
if (led->blue)
   color &= ~(BLUE);
if (led->red)
   color &= ~(RED);
if (led->green)
   color &= ~(GREEN);
retval =
   usb_control_msg(led->udev,
                  usb_sndctrlpipe(led->udev, 0),
                  0x12,
                  0xc8,
                  (0x02 * 0x100) + 0x0a,
                  (0x00 * 0x100) + color,
                  buffer,
                  8,
                  2 * HZ);
kfree(buffer);
```

This function starts out by setting all bits in the variable color to 1. Then, if any LEDs are to be enabled, it turns off only that specific bit. We then send a USB control message to the device to write that color value to the device.

It first seems odd that the tiny buffer variable, which is only 8-bytes long, is created with a call to kmalloc. Why not simply declare it on the stack and skip the overhead of dynamically allocating and then destroying it? This is done because some architectures that run Linux cannot send USB data created on the kernel stack, so all data that is to be sent to a USB device must be created dynamically.

LEDs in Action

With this kernel driver created, built and loaded, when the USB lamp device is plugged in, the driver is bound to it. All USB devices bound to this driver can be found in the sysfs directory for the driver:

```
$ tree /sys/bus/usb/drivers/usbled/
/sys/bus/usb/drivers/usbled/
`-- 4-1.4:1.0 ->
../../../../devices/pci0000:00/0000:00:0d.0/usb4/4-1/4-1.4/4-1.4:1.0
```

The file in that directory is a symlink back to the real location in the sysfs tree for that USB device. If we look into that directory we can see the files the driver has created for the LEDs:

```
$ tree /sys/bus/usb/drivers/usbled/4-1.4:1.0/
/sys/bus/usb/drivers/usbled/4-1.4:1.0/
|-- bAlternateSetting
|-- bInterfaceClass
|-- bInterfaceNumber
|-- bInterfaceProtocol
|-- bInterfaceSubClass
|-- bNumEndpoints
|-- blue
|-- detach_state
|-- green
|-- iInterface
|-- power
|   `-- state
`-- red
```

Then, by writing either 0 or 1 to the blue, green and red files in that directory, the LEDs change color:

```
$ cd /sys/bus/usb/drivers/usbled/4-1.4:1.0/
$ cat green red blue
0
0
0
$ echo 1 > red
[greg@duel 4-1.4:1.0]$ echo 1 > blue
[greg@duel 4-1.4:1.0]$ cat green red blue
0
1
1
```

This produces the color shown in Figure 3.

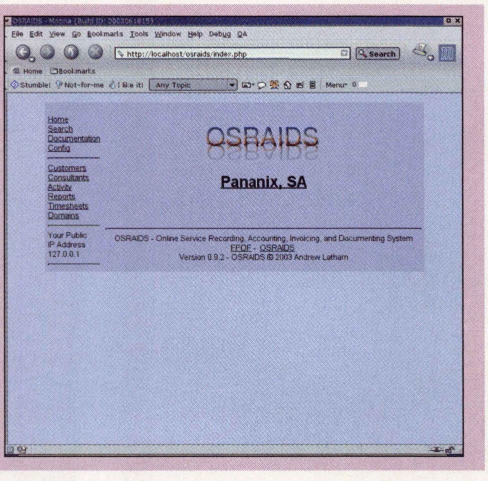
Is There a Better Way?

Now that we have created a simple kernel driver for this device, which can be seen in the 2.6 kernel tree at drivers/usb/misc/usbled.c or on the *Linux Journal* FTP site at (ftp.ssc.com/pub/lj/listings/issue120/7353.tgz), is this really

OSRAIDS—Online Service Recording, Accounting, Invoicing, and Documenting System: lathama.org/osraids

If you have a small service-oriented business and need to keep track of time spent on a project or client and then bill based on time, OSRAIDS may be for you. This is *not*, however, an accounting package, despite the name. This is a time-keeping system that allows you to bill clients, which may be more than adequate for many small service-oriented businesses. You can print PDFs if you have the FPDF package installed. Requires: Web server with PHP support, PHP with MySQL support, MySQL, FPDF (optional) and Web browser.

—DAVID A. BANDEL



http://localhost/osraids/index.php

OSRAIDS

Pananix, SA

Your Public IP Address 127.0.0.1

OSRAIDS - Online Service Recording, Accounting, Invoicing, and Documenting System
FPDF - OSRAIDS
Version 0.9.2 - OSRAIDS © 2003 Andrew Latham

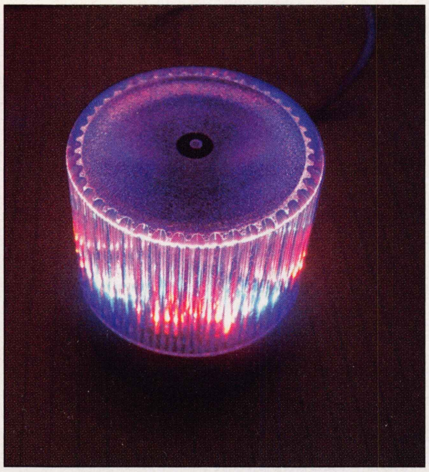


Figure 3. The Device with the Red and Blue LEDs On

the best way to talk to the device? What about using something like usbfs or libusb to control the device from user space without any special device drivers? In my next column, I will show how to do this and provide some shell scripts to control the USB lamp devices plugged in to the system easily.

If you would like to see kernel drivers written for any other types of devices, within reason—I'm not going to try to write an NVIDIA video card driver from scratch—please let me know.

Thanks to Don Marti for bugging me to get this device working on Linux. Without his prodding it would have never gotten finished.

Greg Kroah-Hartman currently is the Linux kernel maintainer for a variety of different driver subsystems. He works for IBM, doing Linux kernel-related things, and can be reached at greg@kroah.com.

rsnapshot: www.rsnapshot.org

I've seen several local/remote backup packages, but this probably is one of the simplest to set up and use. It also makes use of hard links for storage, so the same file isn't stored twice but hard linked, which saves a lot of disk space. Backups can be made as often as you like, and remote backups use SSH. Requires: Perl, rsync and SSH (optional).

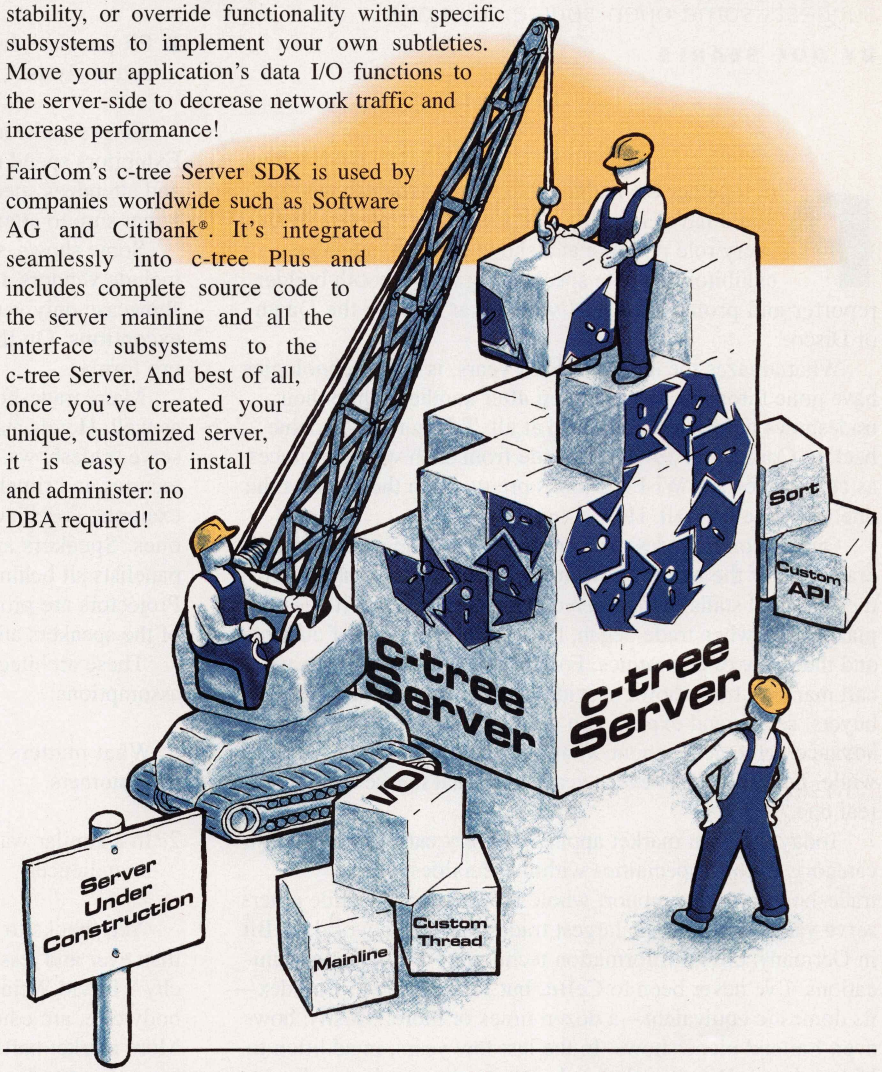
—DAVID A. BANDEL

CUSTOMIZE YOUR DATABASE SERVER WITH THE C-TREE® SERVER SDK

Today's database demands are often too complex for traditional database servers. The functionality and precise level of control you need is simply not available. Perhaps you need alternate sort criteria for your data or a special twist in the threading or communication logic.

FairCom's c-tree® Server SDK allows you to create a customized, industrial-strength server designed for your particular needs. Use FairCom's kernel, with over 20 years of proven stability, or override functionality within specific subsystems to implement your own subtleties. Move your application's data I/O functions to the server-side to decrease network traffic and increase performance!

FairCom's c-tree Server SDK is used by companies worldwide such as Software AG and Citibank®. It's integrated seamlessly into c-tree Plus and includes complete source code to the server mainline and all the interface subsystems to the c-tree Server. And best of all, once you've created your unique, customized server, it is easy to install and administer: no DBA required!



- Enhance our server with your own custom server-side functionality
- Move functionality from the client-side to the server-side to reduce network traffic and increase performance
- Modify or replace entire server subsystems
- Complete source for the server mainline, key server subsystems, and client-side
- Flexible OEM licensing

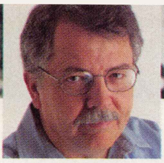
Visit www.faircom.com/ep/lj/sdk today to take control of your server!

FairCom®
www.faircom.com

USA	573.445.6833
EUROPE	+39.035.773.464
JAPAN	+81.59.229.7504
BRAZIL	+55.11.3872.9802

DBMS Since 1979 • 800.234.8180 • info@faircom.com

Other company and product names are registered trademarks or trademarks of their respective owners. © 2002 FairCom Corporation



Showtime

Can we fix tradeshows? Our veteran showman suggests some open-source solutions.

BY DOC SEARLS

Put together the badges I've worn at tradeshows, and you'd have enough to tile a wall. I've played about every role possible at a show: panelist, attendee, exhibitor, keynote speaker, organizer, booth builder, reporter and promoter. And I've been at it since the Dawn of Disco.

What amazes me, after all these years, is how technologies have gone through one revolution after another, while their tradeshows hardly have changed at all. Take a time machine back to Comdex in 1984 and, aside from such wireless graces as cell phones and Wi-Fi, it looks pretty much the same as the one we'll see this fall. How come?

One reason is tradeshows are markets in the ancient and literal sense of the word. They are the industrial equivalents of the tents and stalls that gathered at crossroads and village plazas back when trade began, long before the birth of currency and the study of economics. For the industrial categories we call markets, tradeshows provide a literal marketplace where buyers, sellers and experts can gather to meet, do business and advance agreements about what makes their category worthwhile. For a few days every year the virtual market becomes a real one.

Today the term market applies to every category and subcategory, even to specialties within specialties. Some tradeshows broadly support whole industrial pies, while others serve various slices. The largest tradeshow in the world, CeBit in Germany, covers information technology and telecommunications. I've never been to CeBit, but I have been to Comdex—its domestic equivalent—a dozen times or more. Mostly, however, I attend niche shows. In the last few years, in addition to Linux shows, I've attended and often spoken at shows about identity management, open source, emerging technologies, computing and communications platforms, Macintosh and Mac OS X, Jabber, BSD, Apache, embedded technology, interactivity, new products and technologies, computer product distribution, blogging, peer-to-peer, activist politics, privacy and security, technology impact on people, customer service, government technology and electronic commerce—to name the shows I can remember attending without looking at old calendars.

Although the subjects of these conferences differed widely, and although they variously were called conferences, conventions, shows, events, expos, workshops and -cons of various sorts, they generally held to long-established tradeshow conventions. Let's look at those conventions:

First, tradeshows gather five types of people:

- Exhibitors, including employees staffing booths.
- Speakers, usually keynoters and panelists for breakout sessions.
- Attendees, who visit vendors and listen to speakers.
- Press and analysts, who report on vendors and speakers.
- PR people, who help transmit vendor information to press and analysts.

Shows make money from both exhibitors and attendees. Exhibitors spend many thousands of dollars for booth space, and attendees spend either a small sum to visit exhibitors or a larger sum to attend keynotes and breakout sessions.

Some shows, such as Pop!Tech (www.poptech.org), don't include vendors. Others, such as Demo (www.demo.com), showcase only vendors' new products. But those shows are exceptions. On the whole, the system I described above is pro forma.

Many tradeshow formalities are defaulted by their venues as well. Hotels and convention centers exist in large measure to serve tradeshows. Doing that requires highly formalized architectures and ritualized procedures. Flexible meeting space, for example, is achieved by partitioning larger rooms into smaller ones. Speakers are expected to stand behind podiums, and panelists sit behind tables on stages at the fronts of rooms. Projectors are provided for presentations or video enlargements of the speakers and panelists.

These architectural and procedural defaults make two assumptions:

1. What matters most is helping vendors sell stuff to customers.

2. In a similar way, knowledge flows top-down, from speakers to audiences.

Yet, thanks to the Net, markets today are more connected than ever and less hierarchical. ("Hyperlinks subvert hierarchy", David Weinberger says.) Customers, along with everybody else, are educating themselves and one another, directly. More market participants are in positions to teach as well as learn. Large corporate players are no longer the primary sources of wisdom about the markets they lead. Every market's wisdom today is highly distributed and able to distribute itself any way it pleases: e-mails, personal publications on-line, wikis, IM, IRC and other forms of social computing, as well as through ad hoc meetings in meat space.

So it's no surprise to hear that attrition now is a problem in the tradeshow business and that the networked marketplace is a part of the problem. *Meetings & Conventions*, in the January 2004 issue, says:

> The instances of attendees circumventing traditional registration to book through on-line third-party providers has avalanched, leaving planners increasingly frustrated and making attrition a major issue for associations.

> In 2004, a number of hotels will team up with planners to fight the problem on a united front. “Hotel companies traditionally have stayed on the sidelines and said, ‘We are protected by a contract. It’s the customer’s problem,’” says Joel Pyser, vice president of field sales for Bethesda, Maryland-based Marriott International Inc. “Well, we can’t afford to do that any longer. No one enjoys presenting a customer with an attrition bill. We have to partner with them to solve it.”

In this case, notice, the customer is the company that puts on the tradeshow, not the hotel guest who attends the tradeshow. That individual is still a consumer:

> Web surfing for the lowest rate has become a consumer sport. It also is a major thorn for planners managing room blocks. The Alexandria, Virginia-based International Society of Hospitality Consultants estimates on-line bookings accounted for 30%–40% of all reservations booked within a four-week arrival window in 2003—up 20 percent over pre-9/11 numbers.

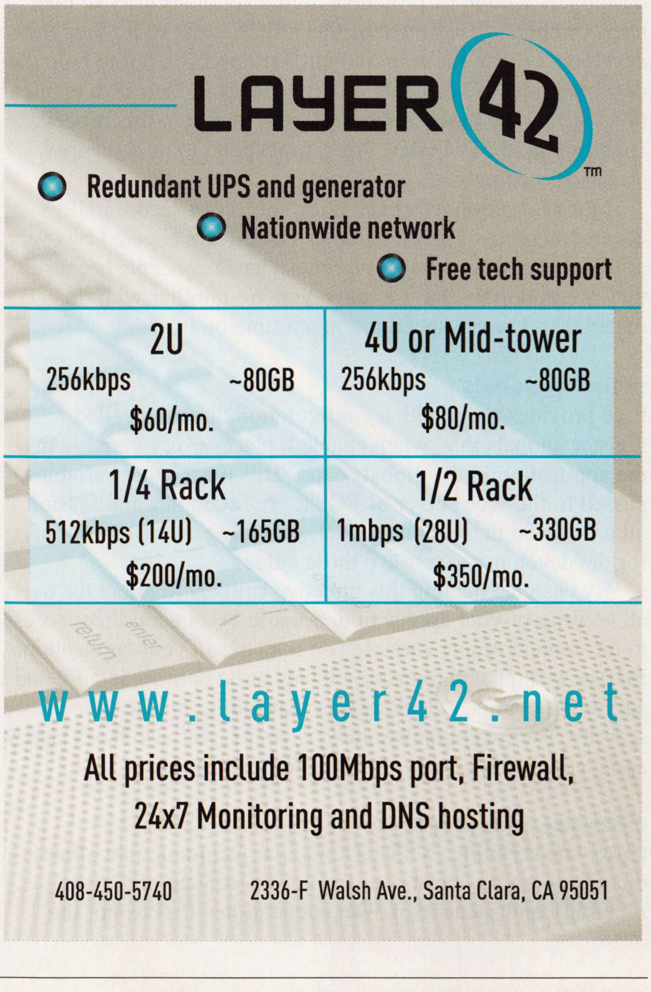
Of course, this top-down perspective will remain ingrained to the degree that the bulk of income flows from large corporate customers. Still, it seems to me that this whole industry could benefit from the same open-source value system and development methods that caused the growth of Linux and the Net.

In that spirit, I’d like to share the best of what I’ve learned from the best of the tradeshows I’ve attended over the last few years. Here goes:

- Hold collegial meetings, not sessions. Some of the best sessions at conferences are the BoFs, or birds-of-a-feather sessions, held after hours. O’Reilly does an excellent job of aspiring to this format, even in its standard sessions.
- Insist on long Q&A periods after lectures, with many working microphones to pass around the audience.
- Record all sessions and make them available on-line in open audio formats. This also helps sell attendees on coming to the next conference.
- Use conversation-friendly venues. Pop!Tech is held in an old opera house in downtown Camden, Maine. The feeling of the whole place is cozy and comfortable, and somehow it helps make the audience more involved with each talk.
- Try forbidding vendor pitches. The first BloggerCon wasn’t a vendor venue. It was held at Harvard, in lecture rooms. But vendor pitches still were discouraged, and it helped open conversation to topics that transcended market economics.
- Make the Web a living and permanent resource and document archive. Here nobody beats O’Reilly’s system, which includes the company’s own writers, as well as pointers to writing done by outsiders covering or attending the event.
- Provide wireless Net connections. It was at Esther Dyson’s PC Forum that I first witnessed a complete power shift from stage to audience, thanks to Net-connected Wi-Fi. That was three years ago. Since then many other events have added Wi-Fi, but the practice is far from standard. It should be.
- Start topical conversations in advance of the show. Jerry Michalski pioneered this when he was involved in PC Forum, many years ago. Today his retreats are meetings for a persistent community that stays in touch by e-mail, wiki and other methods. Countless constructive new conversations have been started and sustained by these off-conference methods.
- Hold hack-a-thons. Your event’s name doesn’t need to end in hack for hack-a-thons to work. ApacheCon launches itself with a hack-a-thon that’s not to be missed, even if you’re not a hacker. It begins the event with a sense that *stuff can be done*.

If you have any other ideas, send them here or to your friendly conference planner. If we do this thing right, we might revolutionize another industry. ■

Doc Searls (doc@ssc.com) is senior editor of *Linux Journal*. His monthly column is Linux for Suits and his biweekly newsletter is SuitWatch.



LAYER 42™

Redundant UPS and generator
Nationwide network
Free tech support

2U
256kbps ~80GB
$60/mo.

4U or Mid-tower
256kbps ~80GB
$80/mo.

1/4 Rack
512kbps (14U) ~165GB
$200/mo.

1/2 Rack
1mbps (28U) ~330GB
$350/mo.

www.layer42.net

All prices include 100Mbps port, Firewall, 24x7 Monitoring and DNS hosting

408-450-5740
2336-F Walsh Ave., Santa Clara, CA 95051

Real-World PHP Security

Understanding the most common security threats to PHP applications is the first step to securing yours. **BY XAVIER SPRIET**

During the past two years, the core PHP developers have done an incredible job of providing the PHP user community with powerful technology that has been able to perform remarkably well in many environments. As Web applications become more popular, Web developers must face an increasing amount of possible security vulnerabilities that have the potential to compromise their work seriously. Many tutorials, books and articles have been published as new techniques are developed. These new emerging threats, however, have not received the focus they deserve.

This article is aimed at professional and open-source PHP developers who must provide a high level of security to their users or clients. The intent of this article is not to provide the developer with a question-and-answer approach but to help the developer identify possible security issues in their own applications during the design process. In the long run, this process enables you, the PHP developer, to respond to new security threats accordingly.

Many articles have covered the subject of secure PHP development, and the same topics usually are covered by every article. Here, I quickly go over those basic concepts because they are important, but I assume you are familiar with this material so I won't spend too much time on it.

register_globals

PHP provides users with a configuration directive called register_globals that, when enabled, places every variable in the application in the global scope. This means that variables passed to the Web server as POST, GET, cookies and session all are placed in the same basket, providing a convenient way for the developer to retrieve those values.

By design, enabling this directive is likely to affect the overall security of your application, because users gain direct access to the content of any variable you may use in your application. PHP now ships with register_globals turned off by default, and I strongly recommend leaving it at that setting for the sake of security. The exception would be if your server also hosts legacy applications that assume this directive is turned on.

Cross-Site Scripting

Cross-site scripting (XSS) is a popular technique that allows the user to gain control over the layout, content and overall reliability and security of Web applications. PHP is not the only technology vulnerable to this technique, mostly because it is not really a flaw in the language. Instead, it is more of a concept pertaining to the design of Web applications in general.

Cross-site scripting exists in many different forms, but a popular method is to inject HTML or JavaScript code in form fields in order to make your application display content that otherwise should not be displayed. This concept illustrates the importance of always filtering any kind of input for your application, whether it comes from a user, another site or even from the database. The PHP function htmlentities() is generally a good way of preventing this type of attack.

GET Variables

Having the ability to provide users with a URL they can use to get back to where they are later on is critical for most Web applications. But as a developer, it is important to be able to determine what information the user should be able to access in any possible way. By manipulating the content of the query string, the user gains the ability to modify the content of the variables used by your application.

Preventing this type of event from happening is more complex than simply filtering the input, but this is still a step in the right direction. What is perhaps the most reliable way to secure your applications against this type of attack is setting up a robust data-flow scheme for your application and a solid error-control system.

SQL Injection

This type of malicious attack on a Web application can have devastating consequences that go beyond the scope of most other attacks, such as cross-site scripting, because it has the potential to destroy your database and its content permanently and completely.

The concept of SQL injection is quite simple. Most Web applications accept parameters as input from POST and GET variables and from cookies. This input often is used inside an SQL query as a parameter, thus providing the user with dynamic content. If the user has any idea of what your database looks like, they technically should be able to alter the parameters you use to inject SQL commands in to your query.

Let's look at a quick example. Your application accepts data from a form as POST. The goal is to display *x* records from the database, where users can modify *x* to fit their needs. Therefore,

HIGH PERFORMANCE COMPUTING
SOLUTIONS FOR YOUR WORLD

HPC
SOLUTIONS

Cluster-tailored Solutions
- Life Sciences (BioBrew)
- Digital Content Creation (RenderBrew)
- Petrochemical (PetroBrew)

Cluster Management Software
- Callident Rx
- Effortless Installation
- Ease of Management
- Peace of Mind

Servers
- 1U, 2U, 4U Options
- 32 and 64-Bit Platforms
- Cost Effective

Storage
- 36 Hot-Swap Drives in a 4U
- Serial ATA Technology
- Scales to Your Needs

Database Solutions
- Topspin Switched Computing
- Oracle RAC 9i and 10g
- SAN Connectivity

promicro®

intel. premier PROVIDER 2003

Call us today at (866) 656-4276 or visit us on the web at www.promicro.com

Intel Inside logo, and Xeon is a registered trademark of Intel Corporation or its subsidiaries in the United States and other countries.
© 2003, Promicro. All rights reserved. Promicro is a registered trademark of Promicro Systems, Inc.

Listing 1. Building an SQL Query in PHP Based on POST Variables

```
<?php
$query = "SELECT id, name FROM `records` LIMIT "
         . $_POST[´NUM´];
$result = $db->select($query);
?>
```

Listing 2. A Malicious Form Used to Perform the SQL Injection Attack

```
<form action="example.com/form.php" method="POST">
<input type="text" name="NUM"
        value="5; DELETE FROM `records`">
<input type="submit">
</form>
```

Listing 3. A Simple "Harmful SQL Commands" Filter

```
<?php
function filter_sql($input) {
    $reg = "(delete)|(update)|(union)|(insert)";
    return(eregi_replace($reg, "", $input));
}
?>
```

your form simply has a field called NUM that provides your script with that value. Listing 1 illustrates this process. In this case, a user could forge an HTML form that would send a carefully crafted value that in turn, would empty your table.

If the user decides to create a form like the one presented in Listing 2, your end result would look like this:

```
SELECT id, name FROM `records` LIMIT 5;
DELETE FROM `records`
```

There obviously are simple ways to counter such attacks, but I have noticed that a large number of applications have no facility to protect themselves from this type of attack.

In our particular example, calling the intval() function to convert NUM to an integer would have provided a decent level of security against SQL injection. However, it is important to understand that developers can't think about every single parameter used in all of their SQL queries. Therefore, what you really need to do is streamline this process in your applications.

Because modern Web-based applications commonly tend to gravitate toward a core module or some kind of centralized switchboard system, it becomes easy to implement such a facility application-wise. The details of the implementation of streamlined facilities for your applications are covered later in the article. For now, take note of the following quick tips that will help you build your own solution:

1. Use regular expressions to filter SQL commands: this method is not appropriate if you intend to accept text from users, but it does a good job of stopping SQL injection by filtering out SQL keywords (Listing 3).

Listing 4. Typical Usage of the Mcrypt Extension

```
<?php
/* Create your key at random
   but keep it handy as you
   will use it to decrypt later
*/
$key = "AOQKJLCLIGAKJHSD
        <NKLXASLUIHJKHAS
        OIUDSgfuyJKLBLKU";

$string = $_POST[´password´];

/* First, you must open the encryption module
   provided by Mcrypt */
$mod = mcrypt_module_open (´blowfish´,´´,´ecb´,´´);

/* You must then create an Initialization Vector
   based on a size and a source.
   Your source can be custom, but some constants
   are available.
   Defining the size of the vector depends on the
   module you are using */
$iv_size = mcrypt_enc_get_iv_size($mod);

/* The initialization vector will be based on $size
   characters from the source /dev/random in our
   example */
$iv = mcrypt_create_iv($iv_size,MCRYPT_DEV_RANDOM);

/* The next step is to ensure that your key is not
   too big and truncate it if necessary */
$max_key_size = mcrypt_enc_get_key_size($mod);
$key = substr($key,0,$max_key_size);

/* You must then initialize the encryption
   mechanism through mcrypt_generic_init */
mcrypt_generic_init ($mod,$key,$iv);

/* You can now encrypt your data through
   the use of mcrypt_generic. The function
   will return your encrypted data */
$encrypted = mcrypt_generic($mod,$string);

/* Once you have finished using Mcrypt, you
   must free the buffers used during the process */
mcrypt_generic_deinit ($mod);

/* Finally, you must close the encryption module
   you have used*/
mcrypt_module_close ($mod);

/* Now here is how we can decrypt our data: */
$padded = // see next line
mcrypt_decrypt(´blowfish´,$key,$encrypted,´ecb´,$iv);
/* At this point, our decrypted string has been
   zero-padded so we need to remove the extra \0s */
$plain = str_replace("\0","",$padded);
echo "Encrypted string: $encrypted<br>";
echo "Decrypted string: $plain<br>";
?>
```

More Specials on the Web 01.04.04

lantechservers.com
866.306.5590

Their Delivery is as FAST as their Servers! Satisfactory A+
— Jeremy O.

5-Years Warranty
All systems are tested and preloaded with your choice of Linux Distribution

LS-1UP240.i

1U IDE Solution

- 1U Lantech Rackmount Chassis
- Single Intel Pentium 4 3.2GHz 800FSB
- 2GB ECC Registered RAM(4GB max)
- 2 x 200GB SATA Hard Drives with 8MB CACH
- 1 x 40GB OS DRIVE (RAID 1/0)
- 1x10/100 Ethernet adapters
- 1x10/100/1000 Ethernet adapters
- 52x CD-ROM Drive Slim
- External 1.44 MB floppy Drive (optional)

Featured at $1,949

LANTECHSERVERS manufactures secure, reliable, innovative, and cpmpetitively priced server tailored to YOUR needs and demands rather then ours. We provide massive computing solutions to various Professional industries such as *Internet DataCenters nationwide, ISP's, Scientific Research facilities, 3D Render Farms, Video/Multimedia Streaming, CAD designers, Video Editing and rendering farms*, and the list goes on and on...**Botton Line, we gurantee your satisfaction! Our specialists are standing by to assist you. Call us today.**

Bring Pluto Miles Closer to Earth

LS-1UX30.i

1U XEON DIGITAL RENDER MAN

- 1U Lantech Rackmount Chassis
- Dual Intel Xeon 3.06GHz - 512k - 533HMz
- **6GB** ECC Registered RAM (12GB max)
- 1 x 40GB DRIVE 7200RPM
- 52x CD-ROM Drive
- 2x10/100/1000 Ethernet adapters
- External 1.44 MB floppy drive (optional)

RENDER MAN

Featured at $3,749

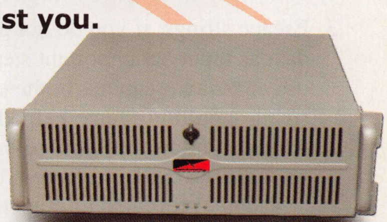
LS-2UX28.s

2U SCSI RAID STORAGE SERVER

- 2U Lantech Rackmount Chassis
- Dual Intel Xeon 2.8GHz - 512k - 533MHz
- 2GB ECC Registered RAM(12GB max)
- 6 x 73.6GB HotSwap SCSI Hard Drives
- 2 x 20GB OS DRIVE -RAID 0/1/5/10
- 2x10/100/1000 Ethernet adapters

440GB SCSI

Featured at $4,995



LS-4UX30.si

4U IDE Storage MAX

- 4U Lantech Rackmount Chassis
- Dual Intel Xeon 3.06GHz 533FSB
- 2GB ECC Registered RAM(12GB max)
- 10 x 250GB SATA Hard Drives 8MB
- 2 x 20GB OS DRIVE -RAID 0/1/5/10
- 2x10/100/1000 Ethernet adapters

2.5TB SATA

Featured at $5,995

ISO 9001 Certified

Lantechservers.com is not responsible for any typographical errors. Pricing and availability are subject to change without notice.

Listing 5. Error Reporting through Assertions

```
<?php

/* You can toggle assertions throughout your entire
   application by switching ASSERT_ACTIVE to 1 or 0
*/
assert_options(ASSERT_ACTIVE,1);

/* We do want the application to exit when an
   assertion fails. (in this example)
*/
assert_options(ASSERT_BAIL,1);

/* In our example, we will do the error reporting
   ourselves so we turn off the default warnings
*/
assert_options(ASSERT_WARNING,0);

/* display_error will be the name of our custom
   function that will be called if an assertion
   fails
*/
assert_options(ASSERT_CALLBACK, "display_error");

$email = strtolower($_POST[´email´]);
$parts = array();

// Building your regular expression
$regex = "^([.\´a-z0-9]+)@([.\´a-z0-9]+)$";

/* Checking for valid format and splitting
   the email address at the same time
   Note the special formatting. Everything
   is in quotation marks and the error is
   commented. We will extract this error
   later through regular expressions.
*/
assert("ereg(\$regex, \$email, \$parts); /*
       Invalid email address: $email */");

/* This block will not be executed if the
   assertion fails so we can safely go on */
$username = $parts[1];
echo "Welcome home, " . $username;

// This is our ASSERT_CALLBACK function
function display_error($file, $line, $error) {

    // This block will extract the comment message
    $regex = "(.*)/\* (.*)\*/";
    $parts = array();
    ereg($regex, $error, $parts);
    $msg = $parts[2];

    // And we can output a nice little report
    echo "
    <table bgcolor=\"#bbbbee\">
    <tr><td colspan=´2´ align=´center´>
    <b>Error Report</b>
    </td></tr>
    <tr><td>File:</td><td>$file</td></tr>
    <tr><td>Line:</td><td>$line</td></tr>
    <tr><td>Message:</td><td>$msg</td></tr>
    ";

}

?>
```

2. Use assertions: assertions are covered in more detail later in this article.

3. Escape strings: if you do not expect to be accepting binary data as input, an important step in securing your input is the use of string escaping. In the example above, escaping the string would not have helped, however; many SQL injection attacks are based on exiting the SQL query prematurely and injecting a new query inside. This is efficiently prevented through the use of functions, such as mysql_escape_string().

Encryption

Sensitive information often is stored on database servers and other storage facilities for later retrieval. At this point, it is critical to have at your disposal a facility that allows you, as a developer, to secure that data at storage time and retrieve the information you are looking for when you need it.

PHP offers an extension that allows developers to use the Mcrypt Library (mcrypt.sf.net) to secure data by encrypting it and later decrypting it. The documentation of the Mcrypt extension for PHP is located at www.php.net/mcrypt, and it should be studied carefully before implementation.

The Mcrypt extension supports an impressive array of algorithms, including triple-DES, Blowfish, Twofish and Two-Way. Using the Mcrypt extension is not a very intuitive process if you are not familiar with encryption; it can become quite confusing because of the variety of block algorithms and encryption modes available. Refer to Listing 4 for a sample of what the Mcrypt extension offers and how to use it.

Assertions

Assertions provide the PHP developer with a way to implement error control and preserve the integrity of data. This is not a security-related feature of PHP, and it is implemented in many mainstream languages, such as C or Python, so why am I bringing it up now? Simply put, error control is the first step in providing efficient security for your users or your clients.

Assertions are implemented in PHP through the use of two functions, assert_options() and assert(). The former should be called in your application's initialization or configuration file, and the latter should be implemented anywhere in your code where you need to enforce the validity of your input. Listing 5 demonstrates how assertions can be used to create an error-control system that generates a simple report when an assertion fails. The PHPUnit Project is a complete unit testing suite freely technology you can count on

64-bit!

Itanium® 2 and Opteron™ Clusters

Robust Solutions for Demanding HPC Applications

Itanium2, Opteron, Xeon, Athlon and Alpha Clusters

Expert Integration...

Microway offers innovative, competitively priced, custom designed clusters with NodeWatch™ management and monitoring tools. Our solutions incorporate the latest processors, proprietary cooling and storage solutions, plus high-speed Myrinet and InfiniBand interconnects for both fine and coarse grain parallel applications.

Superior Service and Tech Support...

We understand that on-time delivery, out-of-the-box reliability and excellent ongoing technical support are critical to our users. Microway offers professional services from specialists with a wide range of expertise in HPC applications. On-site installations and training are also available.

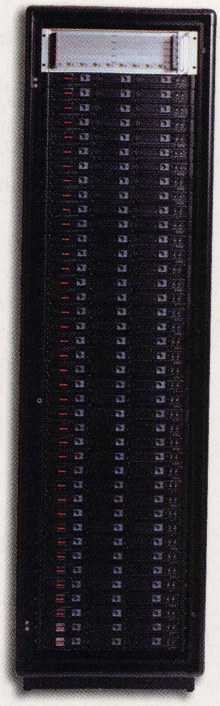
Satisfied Customers...

AT&T, Cessna, GE, GSK, Johnson & Johnson, LANL, LLNL, MBL, Millennium Pharmaceuticals, NIH, Northrop Grumman, Raytheon, Sandia, Seagate, US Air Force, Army, Navy, NASA, NOAA and hundreds of leading universities are among our satisfied customers since 1982.

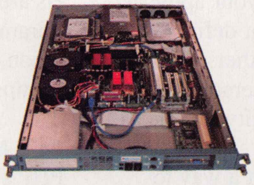
> "The Brain Imaging Research Center (a joint center of Carnegie Mellon University and University of Pittsburgh) decided to purchase our Linux cluster from Microway because of the proven performance of their clusters at Carnegie Mellon in the processing of high-volume brain imaging data. Microway was flexible and helpful at all stages, starting from the initial custom configuration and ending with timely delivery and full installation."
>
> *– Marcel Just, Co-Director, Brain Imaging Research Center*

Call us first at ***508-746-7341***
for quotations and benchmarking services.
Find technical information, testimonials, and
online newsletter at ***www.microway.com.***

See us at **SC2003**, Booth 2227
Phoenix – November 17-20



CoolRak™ Cabinet
Features 3 10" 535 CFM Fans, yielding 7 degrees cooler ambient per node. Available only on Microway 1U Opteron, Athlon and Xeon chassis.

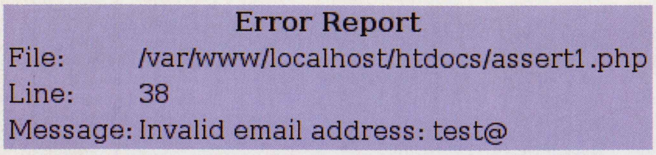


Microway 1U CoolRak™
features dual 2.8GHz Xeons with proprietary Across the Board™ Cooling

Innovative Technology and Services at Competitive Prices

- **Xeon XBlade™ Technology by Microway** Incorporates two dual Xeon XBlades in our 1U chassis for 880 Gigaflops performance per 44U cabinet. Each Xeon XBlade provides two 133 MHz PCI-X sockets enabling next generation interconnects including InfiniBand.
- **NodeWatch™** provides automatic remote monitoring of vital cluster parameters and failsafe shutdown of the cluster. NodeWatch monitors temperatures, voltages, and chassis fans, runs off the master, and can be controlled by a secure web-based GUI. Available on Microway 1U Opteron, Xeon and Athlon chassis.
- **Integrated solutions powered by Platform Computing** Workload Management and Grid Computing software solutions to optimize computing resources. From Microway's Clusterware-powered CW Series, to integration of LSF 5 on any Microway cluster, all Platform solutions are "grid ready" for future system enhancements.
- **Integrated professional MPI/Pro®** enhances cluster value with middleware for high performance, scalable and robust message passing.
- **Fully redundant, highly-available storage systems** based on 2Gbit fiber channel technology for multi-terabyte storage requirements. State-of-the-art storage directors for full access to storage from any cluster node. Easily scales to address rapidly expanding storage requirements.
- **Consulting Services:** Microway offers expertise in designing custom software for financial analysis, genomics, proteomics, plus data warehousing, project management tools, network/data center design, and lab integration/management systems.

Microway®
Technology you can count on℠



Error Report
File: /var/www/localhost/htdocs/assert1.php
Line: 38
Message: Invalid email address: test@

Figure 1. A Sample Report Generated by Listing 5

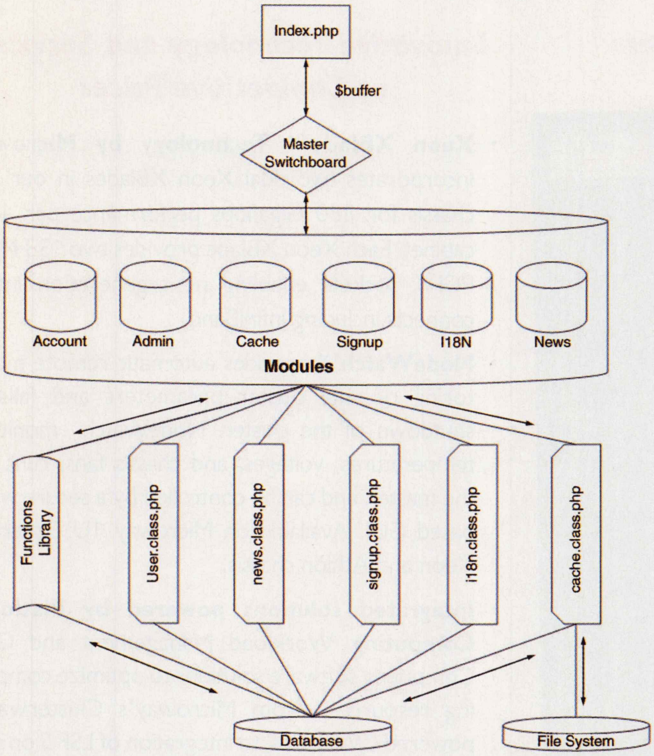


Figure 2. Application Core

available to PHP developers and is based on what we have just done. The PHPUnit's home page is located at phpunit.sf.net.

Data Flow

If you have worked on many different Web projects, chances are you have started using a common structure upon which to base your new projects or you have developed your own. There are many ways to centralize data management in your application, and depending on the set of requirements that define your project, some models are more appropriate than others. In the next few paragraphs, I introduce a simple design template that gives the developer a sufficient amount of scalability and flexibility for most enterprise-grade projects.

What you need to do at this point is implement a way to centralize all your input and force it to go through a filtering facility. Doing so gives you the simplicity you need to implement additional functionality in a modular fashion. In our example, we use the following file hierarchy:

- /index.php: only file in root.
- /lib: libraries, protected by .htaccess.
- /lib/config.inc.php: configuration file.
- /tpl: templates, protected by .htaccess.
- /doc: project and APIs documentation.
- /images.
- /classes: classes, protected by .htaccess.

As illustrated in Figure 2, your application's core is the index.php file, and it has direct access to any library, template, class or configuration file, but the user never has access to those files.

Let's follow, step-by-step, the design illustrated in Figure 2 by taking the example of a user logging in to the application.

1. The user queries index.php with no parameters. Index creates a buffer and passes it over to the switchboard that calls the default module. This module uses a template to display the default page of the application.

2. The user fills in the authentication form and submits the form. The form redirects its output to something like `?module=account&action=login`. The switchboard calls the login function of the account module, which is simply an interface to the user class. The function instantiates an object of the user class. This object is an interface between your module and the database, and it performs the query.

3. The data is sent back from the database to the object and from the object to the module, which in turn, sets up the appropriate session variables, calls the proper template and uses it to modify the buffer. It then sends the response message to the index.

The data flow in this particular model may seem a little confusing at first, but it really is simple. User input is passed quickly to the appropriate module, and error control is implemented on the switchboard level. Other types of inputs are database access and filesystem access, and they are filtered by their appropriate classes. Every class extends a special skeleton class that provides the input filtering facility, so none of the classes have to worry about this.

This model is efficient as it provides a scalable and robust architecture, but keep in mind that many other interesting models are available. For example, you may want to look at the Phrame Project (phrame.sf.net), which provides an implementation of the Model2 approach, a derivative of MVC (ootips.org/mvc-pattern.html).

Safe Mode

PHP's Safe mode is something you should learn to work with whether you are a PHP developer or a system administrator. Safe mode is a set of configuration options that allow the system administrator to alter the behavior of the PHP interpreter by implementing security measures. From a system administrator's point of view, this means you must learn how to implement this feature properly, without making it impossible for developers to set up their applications on your server. From a developer's point of view, you must learn what possibly could get broken in your application if this feature is turned on.

Turning safe_mode on makes sense if you manage a shared

server that serves PHP applications and the PHP developers using this server are not trusted. Enabling safe_mode in your php.ini file effectively makes any file-related operation in any of your scripts impossible unless the UID of the owner of the file is the same as the UID of the running script. PHP also gives you the ability to change this policy while safe_mode is on by turning on the safe_mode_gid option. In this case, PHP checks for the GID of the files you are trying to work with instead of their UID.

It also is good practice to not let your users execute any system binary they want; safe_mode_exec_dir comes into play here. This priceless feature lets you tell PHP not to perform any binary execution, through exec() or any other function, unless the binary is located in the safe_mode_exec_dir, such as /usr/local/php/bin.

Once you have familiarized yourself with the restrictions implemented by PHP when safe_mode is enabled, you should be able to develop software that doesn't break when it's run on servers with this directive enabled. Many ISPs use safe_mode. The simple guidelines to follow are:

- Try to limit file operations, whether read or write, to the files you have provided with your application.
- Do not rely on external software to be installed or executable by your script unless your project is running on only your servers.

System administrators also have at their disposal other powerful tools to ensure the overall security of their systems. These tools include disable_functions that prevent specified functions from being called, as well as options such as open_basedir, which limit any file operation to a specific directory.

The PHP documentation team has provided an extensive amount of literature on the subject. They also have provided documentation for every aspect of safe_mode and related functions and directives.

Xavier Spriet has been developing software in PHP for the past four years. He is the lead developer at eliquidMEDIA International. You can reach Xavier at xavier@wuug.org.

RESOURCES

Mcrypt Extension: **php.net/mcrypt**

Mcrypt Project: **mcrypt.sf.net**

The MVC Paradigm: **ootips.org/mvc-pattern.html**

PHP Documentation: **php.net/manual/en**

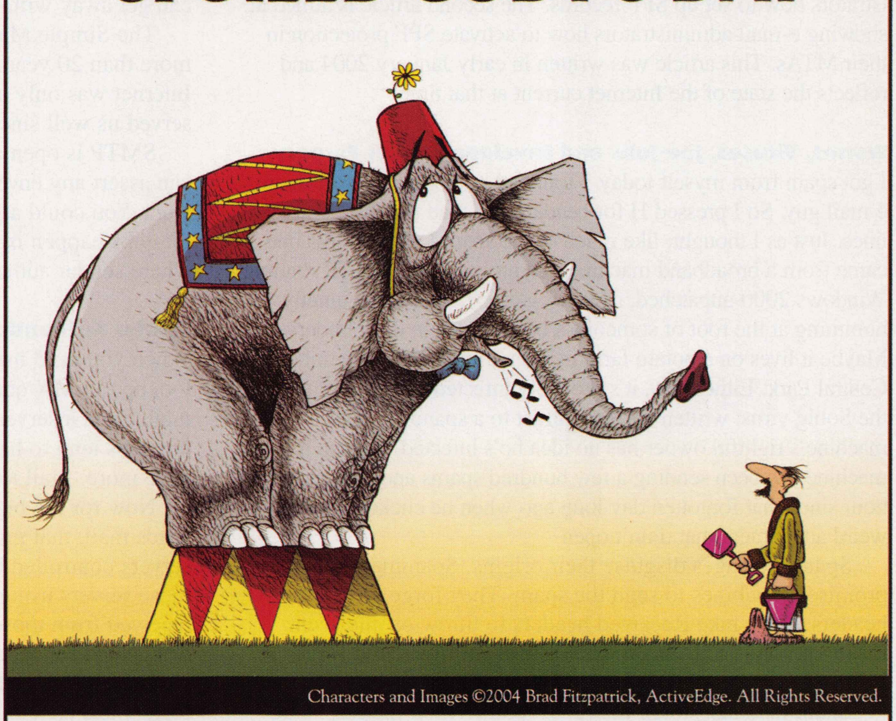
PHP Security: **php.net/manual/en/security.index.php**

The PHPUnit Project: **phpunit.sf.net**

The Phrame Project: **phrame.sf.net**

Safe Mode: **php.net/manal/en/features.safe-mode.php**

ONCE AGAIN,
HEAP PROBLEMS HAD
SPOILED CODY'S DAY



Characters and Images ©2004 Brad Fitzpatrick, ActiveEdge. All Rights Reserved.

Debugging heap allocation problems can be a real chore, but TotalView now has built-in memory features that track memory usage for all processes and can even stop execution at the point that a memory problem occurs. And it's all integrated, so there's no need to interrupt your debug session to invoke an external memory tool. Etnus TotalView is also the best threads debugger available and offers superior C++ support. So, don't forget to download a free fully functional trial of TotalView today.

Try TotalView FREE at www.etnus.com

TotalView, the Most Advanced Debugger on Linux and UNIX

ETNUS
TOTALVIEW

SPF Overview

You can help eliminate the spam problem by making it easy to detect forgeries. Protect your e-mail address reputation with a simple DNS technique.

BY MENG WENG WONG

SPF is an emerging antiforgery standard that aims to prevent worms, viruses and spam from forging arbitrary e-mail addresses as the envelope sender in SMTP. SPF has two parts: domain administrators need to publish SPF records in the DNS, and e-mail administrators need to install SPF-enabled MTAs to read those records. SPF records indicate the servers from which a domain sends outbound mail. Mail coming from anywhere else is considered forged.

This article, the first of a two-part series, explains the concepts and trade-offs involved in SPF protection and shows DNS administrators how to set up SPF records. The second article is aimed at showing e-mail administrators how to activate SPF protection in their MTAs. This article was written in early January 2004 and reflects the state of the Internet current at that time.

Worms, Viruses, Joe-Jobs and Envelope Sender Forgery

I got spam from myself today. I founded pobox.com, and I'm an e-mail guy. So I pressed H for headers and read the Received lines. Just as I thought: like much of the spam I receive, this one came from a broadband machine. It's probably an old PIII running Windows 2000 unpatched, used for gaming and MP3s, quietly humming at the foot of someone's bed, draped in dirty underwear. Maybe it lives on a potato farm in Idaho; maybe it looks out over Central Park. Either way, it's probably infected with a variant of the Sobig virus, written under contract to a spammer. The machine's rightful owner has no idea he's infected, no idea his machine has been sending a few hundred spams and viruses every hour since that forgotten day long ago when he clicked on that weird attachment that didn't open.

Spam messages disguise their origins. Spammers use compromised machines to send the spam. They forge message headers. They fake Received headers to throw off the scent, make up bogus Subjects to trick Bayesian filters and forge From lines pretending to be PayPal or eBay.

Spammers also forge the return path. When messages are undeliverable, they bounce back to the sender whose address is in the return path. Not the From: address in the message headers, but the return path of the SMTP envelope, the RFC2821 MAIL FROM. Often, spamware uses lists of old addresses, or they simply guess common user names or launch a dictionary attack. The result is a lot of bad addresses and a lot of bounces.

Spammers don't want those bounces. They'd rather somebody else receive them. So, they pick an address at random or use the recipient's address. That's how they made it look like I got spam from myself. Sometimes they choose a hated enemy and maliciously forge his address so he gets flooded with thousands of bounces.

In 1997, a spammer forged a return address at joes.com, which then was flooded by so many bounce messages it went down for ten days—and gave the world the term *joe-job*. Hotmail and AOL get joe-jobbed every day: a lot of spam pretends to be from AOL but doesn't really come through their servers. Under conventional SMTP, AOL can't do anything about it. If you put the AOL logo on a T-shirt and tried to sell that shirt, AOL's lawyers would have you ceasing-and-desisting in a heartbeat. But spammers forge @aol.com every day. They can get away with it because they use SMTP.

The Simple Mail Transfer Protocol (SMTP) was designed more than 20 years ago—a kindlier, gentler time. The entire Internet was only a handful of research institutions. SMTP has served us well since then, but it's beginning to show its age.

SMTP is open and trusting. Its rules are relatively lax. You can assert any envelope sender and make up all the headers you want. You could argue today, though, that a protocol that lets joe-jobs happen is a little too open, a little too trusting. That's where sender authentication comes in. SPF tightens the rules.

Sender Authentication with SPF

When you send mail to a domain, your MTA does a DNS lookup (an MX query) to find out to which server to route the mail. Such a server is called a mail exchanger (MX). Small domains tend to have only one MX server. Big domains tend to have more. Mail *to* a domain goes *to* its MX servers.

Now for the big idea. In 99% of all cases, when a domain sends mail, that mail originates from a relatively small set of servers controlled by that domain. The domain could designate those servers using the DNS, then announce that any mail not received from those servers probably is forged. That's called a designated sender scheme (Figure 1).

Designated sender schemes are useful because they help fight forgery and are easy to set up. After all, domain owners already know which servers send mail from that domain. When I say send mail from that domain, I mean originate an SMTP transaction where the MAIL-FROM envelope sender shows that domain. I'm not talking about the From: header. This is an important distinction.

Mail from a domain tends to come from a small number of servers. That's true for domains large and small. Mail from aol.com comes from AOL's servers. Mail from my personal domain comes from my personal servers. It certainly doesn't come from a machine covered in dirty underwear.

Many ISPs already are implementing these kinds of rules in a haphazard and often slightly broken way. The problem is, one

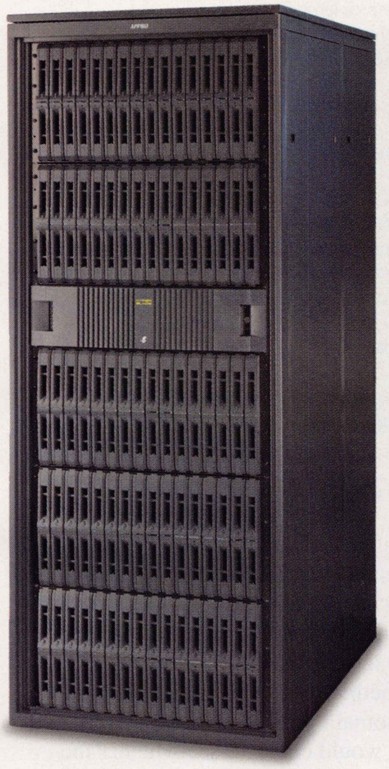
High-Performance Server & Cluster Solutions

Appro servers based on AMD Opteron™ processors boost performance on existing Windows® and Linux applications.

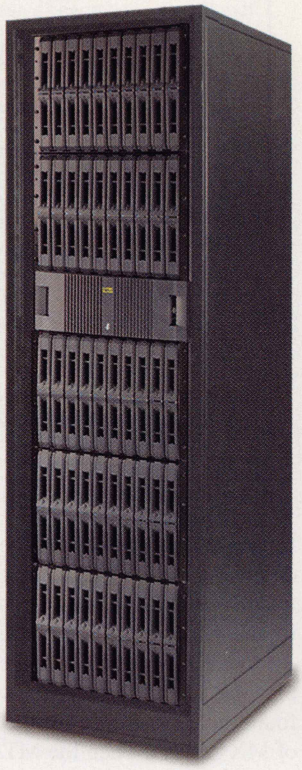
Double the Density!

Appro *Hyper*Blade Fully Integrated Cluster Solutions

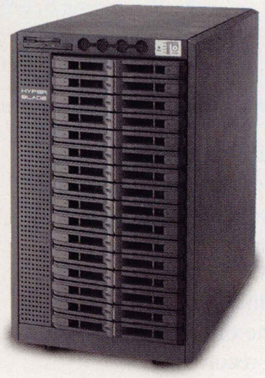
Full-Cluster
Up to **80** nodes



Mid-Cluster
Up to **50** nodes



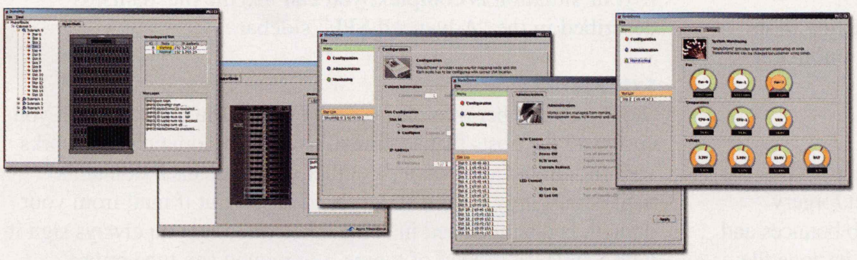
Mini-Cluster
Up to **17** nodes



- Flexible, modular and scalable architecture for rapidly changing environment
- Blade servers fit all cabinet sizes including a standard 19" rack offering the ability to mix and match cluster building blocks
- Dual AMD Opteron™ processors per node
- Up to 16GB ECC DDR Memory per node
- One PCI-X slot per node
- Up to two internal HDDs
- Remote Management Options
- High-Speed Interconnect Options: Myrinet, Dolphin, Quadrics and Infiniband

Appro BladeDome Remote Server Management

✔ *Specially Created for the Appro* HyperBlade *Servers*
✔ *Simple and Cost-Effective Solution to System Administrators*



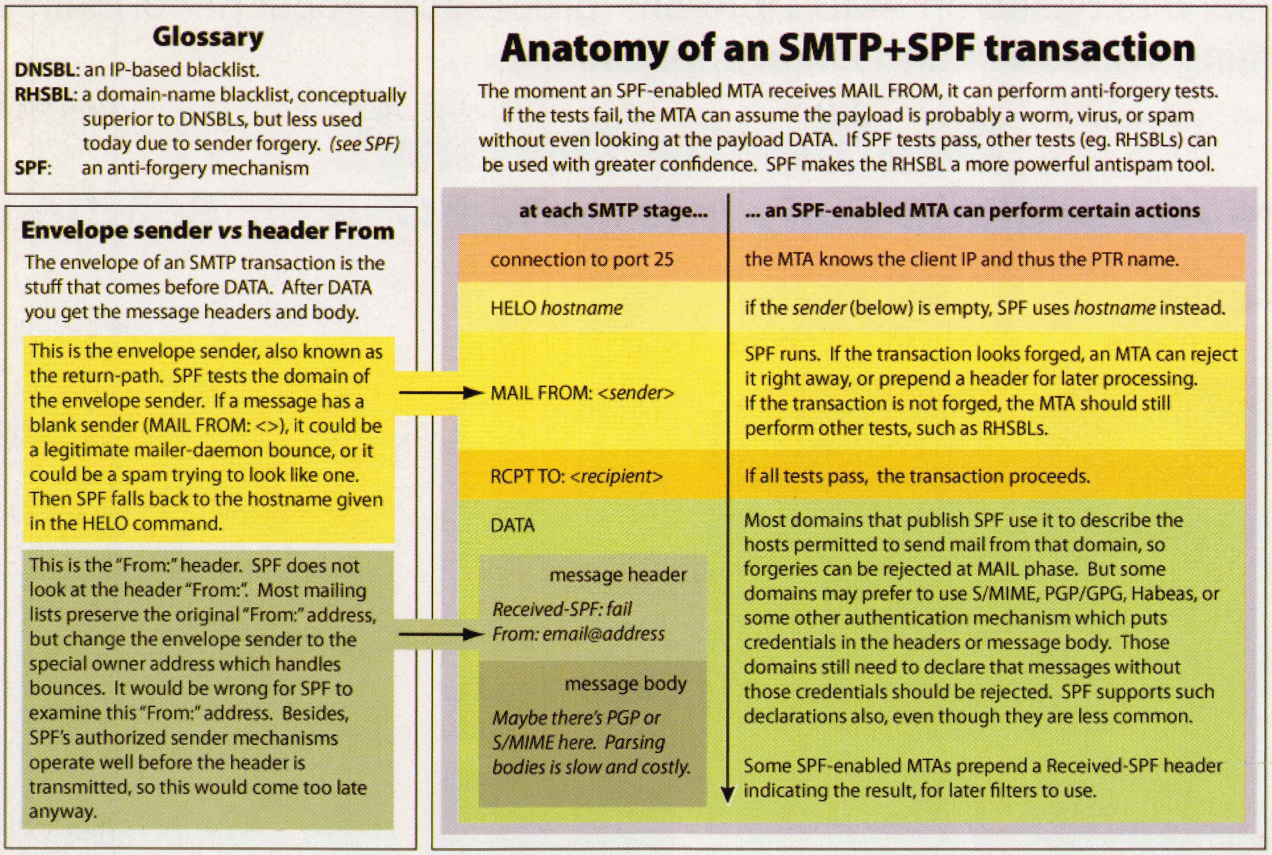
- GUI & command line interfaces
- In-band and out-of-band control
- Remote reset and power cycle
- Platform Monitoring:
  fan fail, over-temperature & voltage
- Failure alert e-mail notifications
- Enhaced Security
- Multiple user account set-up
- BCC Redundancy

AMD Opteron™ Processors - Shorten run-time cycles and increase bandwidth for processing computing requests.
- 32 bit applications while you migrate to 64 bit computing for long-term investment protection.

APPRO
HPC Cluster Solutions

For more information, please visit **www.appro.com**
or call Appro Sales at **800.927.5464**, **408.941.8100**.

AMD 64 Opteron



Glossary

DNSBL: an IP-based blacklist.
RHSBL: a domain-name blacklist, conceptually superior to DNSBLs, but less used today due to sender forgery. *(see SPF)*
SPF: an anti-forgery mechanism

Envelope sender *vs* header From

The envelope of an SMTP transaction is the stuff that comes before DATA. After DATA you get the message headers and body.

This is the envelope sender, also known as the return-path. SPF tests the domain of the envelope sender. If a message has a blank sender (MAIL FROM: <>), it could be a legitimate mailer-daemon bounce, or it could be a spam trying to look like one. Then SPF falls back to the hostname given in the HELO command.

This is the "From:" header. SPF does not look at the header "From:". Most mailing lists preserve the original "From:" address, but change the envelope sender to the special owner address which handles bounces. It would be wrong for SPF to examine this "From:" address. Besides, SPF's authorized sender mechanisms operate well before the header is transmitted, so this would come too late anyway.

Anatomy of an SMTP+SPF transaction

The moment an SPF-enabled MTA receives MAIL FROM, it can perform anti-forgery tests. If the tests fail, the MTA can assume the payload is probably a worm, virus, or spam without even looking at the payload DATA. If SPF tests pass, other tests (eg. RHSBLs) can be used with greater confidence. SPF makes the RHSBL a more powerful antispam tool.

at each SMTP stage...	**... an SPF-enabled MTA can perform certain actions**
connection to port 25	the MTA knows the client IP and thus the PTR name.
HELO *hostname*	if the *sender* (below) is empty, SPF uses *hostname* instead.
MAIL FROM: <*sender*>	SPF runs. If the transaction looks forged, an MTA can reject it right away, or prepend a header for later processing. If the transaction is not forged, the MTA should still perform other tests, such as RHSBLs.
RCPT TO: <*recipient*>	If all tests pass, the transaction proceeds.
DATA message header *Received-SPF: fail* *From: email@address* message body *Maybe there's PGP or S/MIME here. Parsing bodies is slow and costly.*	Most domains that publish SPF use it to describe the hosts permitted to send mail from that domain, so forgeries can be rejected at MAIL phase. But some domains may prefer to use S/MIME, PGP/GPG, Habeas, or some other authentication mechanism which puts credentials in the headers or message body. Those domains still need to declare that messages without those credentials should be rejected. SPF supports such declarations also, even though they are less common. Some SPF-enabled MTAs prepend a Received-SPF header indicating the result, for later filters to use.

Figure 1. With SPF, one mail server can check whether another server really is associated with the address the mail claims to be from.

ISP doesn't know the insides of another ISP, and it's easy to guess wrong. Maybe aol.com's mail servers also originate mail for aol.net or vice versa. Wouldn't it be better if AOL themselves announced their designated servers in a simple, flexible, extensible, open format that everybody could use?

Well, they do. SPF is a standard, flexible, extensible, open format that everybody can use. At the time of this writing, AOL recently had started publishing their SPF record.

MTAs can interpret that record and use it to tell whether mail that claims to be from @aol.com is a fake.

All this rule tightening is purely voluntary: domains that don't publish SPF records can continue to send mail as before. Some unusual domains might be served better by not publishing SPF; it's up to them. But most domains should want to use SPF.

To publish SPF, a domain has to add only one line to its zone file. That line is a TXT record, and you can publish it today. Let's see what the TXT record looks like.

SPF by Example

Suppose example.com wants to publish SPF. It expects MTAs everywhere to read its SPF record and use it to reject forgery attempts. It hopes SPF reduces the volume of joe-job bounces and bogus abuse reports. So it adds the following line to its zone file:

```
example.com.  IN TXT  "v=spf1 a mx ptr -all"
```

The v=spf1 version string identifies this as an SPF record. The -all means reject all mail by default. Domains that don't send any mail, such as altavista.com, can get by with simply `v=spf1 -all`. But if the domain does send mail, it declares mechanisms that describe how legitimate mail should look. Mechanisms go in the middle, before -all. The first mechanism to match provides a result for the SPF query. `-all` always matches and so belongs at the end.

Mechanisms are interpreted left-to-right. Using `v=spf1 a mx ptr -all` first would check whether the connecting client was found in the A record for the domain or, failing that, in its list of MX servers. Then the MTA would check to see whether the hostname of the client matched the domain. If none of the mechanisms matched, -all would be evaluated, the result would be fail and the MTA would be justified in rejecting the mail.

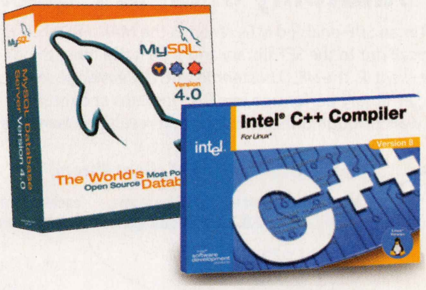
A, MX, PTR and IP4 are enough for the overwhelming majority of domains. The setup wizard at spf.pobox.com/wizard.html can help you configure SPF for your domain. But if your situation is complex, you can use the mechanisms described in the "Advanced SPF" sidebar.

Extensibility

SPF has a number of built-in mechanisms. The basic ones let you designate the hosts that send mail from your domain. This works well for almost all domains out there, because each domain's mail comes only from a small set of hosts. But if mail from your domain is distinguished in some other way, say you always sign it with S/MIME, instead of typing `a` or `mx` you can type `smime`.

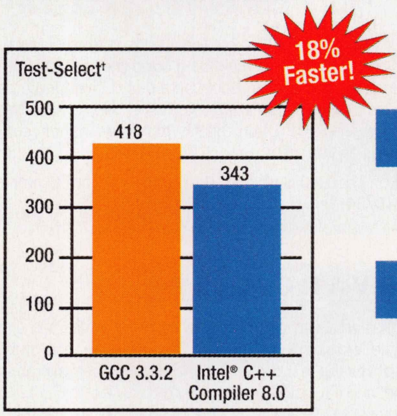
Using designated sender mechanisms (A, MX, PTR and IP4) is one possible approach to sender authentication. New sender authentication methods are being developed. SPF is extensible, though, so it can work gracefully with them. SPF plugins that understand future extension mechanisms will be able to interpret them correctly. SPF plugins that don't understand those mecha-

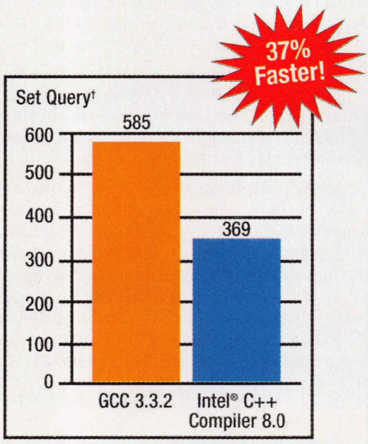
Maximize Your Application Performance



Intel® C++ Compiler 8.0 and MySQL®

The new Intel C++ Compiler 8.0 for Linux achieved the impossible: it made MySQL—one of the fastest databases in the world—even faster!





† Intel® Pentium® 4, 2.8 GHz, 512KB Cache, 256MB RAM, SuSE 8.2, MySQL 4.1.1

Performance

Outstanding performance on Intel architecture including Intel® Pentium® 4, Intel® Xeon™ and Intel® Itanium® 2 processors.

Compatibility

Windows

- Integrates into Microsoft Visual Studio* .NET
- Source/binary compatible with MS Visual C++

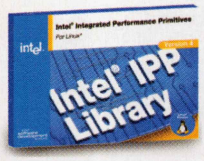
Linux

- Source/binary compatibility with GCC 3.3

Support

1 year of product upgrades and Intel Premier Support included

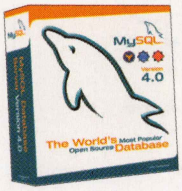
> *"The MySQL database is designed from the start to maximize speed. The new Intel C++ Compiler 8.0 amplified this native speed with proven performance gains of more than 20 percent over using GCC 3.3, extending MySQL's position as one of the fastest, most popular databases in the world."*
>
> *—David Axmark*
> *MySQL AB Co-Founder and Vice President*



Intel® Integrated Performance Primitives Library 4.0

NEW!

Highly optimized code for graphics, multimedia, math and signal processing.



MySQL® Database Server

The world's most popular open source database!

YOU SAVE UP TO $20!	Paradise #	Retail	Discount
MySQL, Box w/GPL License & Manual	MQL 0100	$69.95	$59.99
Intel® C++ Compiler 8.0 for Windows	I23 0B40	$399.00	$378.99
Intel® C++ Compiler 8.0 for Linux	I23 0B41	$399.00	$378.99
Intel® IPP 4.0 for Linux	I23 0B46	$199.00	$189.99

FREE trials available at:
programmersparadise.com/intel

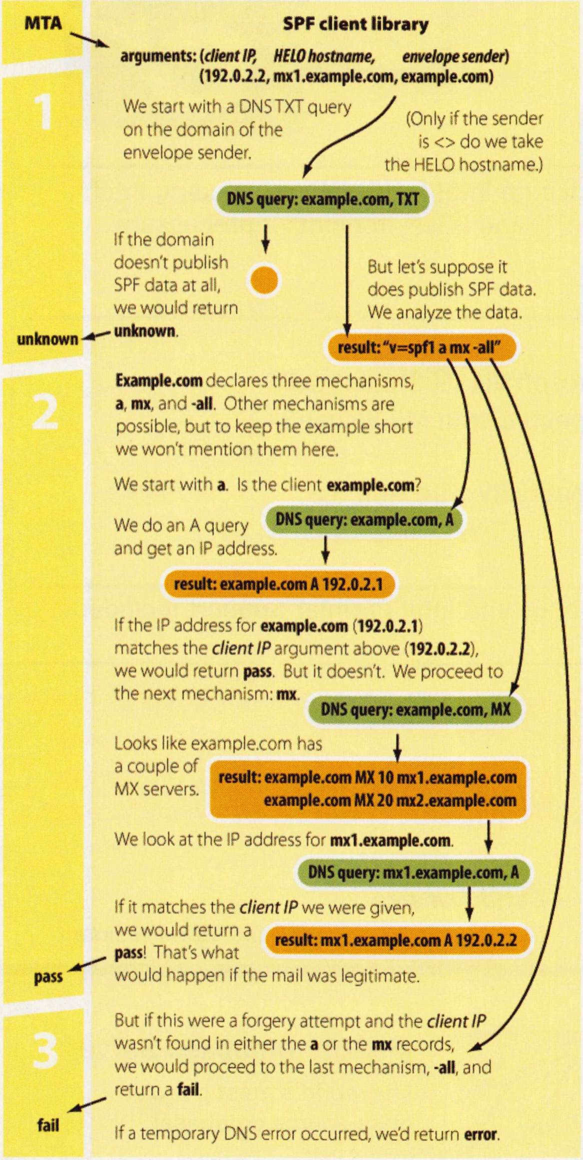
Programmer's Paradise®

To order or request additional information call:
800-423-9990
Email: intel@programmers.com

© 2004 Intel Corporation Intel, the Intel logo, Pentium, Itanium, Intel Xeon and VTune are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries. *Other names and brands may be claimed as the property of others.

Anatomy of an SPF query

After an SPF-enabled MTA receives the MAIL FROM command, it can call out to the SPF library and pass it the three things it needs: the *client IP*, the *HELO hostname*, and the *envelope sender* address. The SPF library performs a few DNS lookups and returns a response. We'll walk through three scenarios that result in **unknown**, **pass**, and **fail**.


MTA
SPF client library
arguments: (*client IP*, *HELO hostname*, *envelope sender*)
(192.0.2.2, mx1.example.com, example.com)
1
DNS query: example.com, TXT
unknown
result: "v=spf1 a mx -all"
2
DNS query: example.com, A
result: example.com A 192.0.2.1
DNS query: example.com, MX
result: example.com MX 10 mx1.example.com
example.com MX 20 mx2.example.com
DNS query: mx1.example.com, A
result: mx1.example.com A 192.0.2.2
pass
3
fail

We start with a DNS TXT query on the domain of the envelope sender.

(Only if the sender is <> do we take the HELO hostname.)

If the domain doesn't publish SPF data at all, we would return **unknown**.

But let's suppose it does publish SPF data. We analyze the data.

Example.com declares three mechanisms, **a**, **mx**, and **-all**. Other mechanisms are possible, but to keep the example short we won't mention them here.

We start with **a**. Is the client **example.com**?

We do an A query and get an IP address.

If the IP address for **example.com** (**192.0.2.1**) matches the *client IP* argument above (**192.0.2.2**), we would return **pass**. But it doesn't. We proceed to the next mechanism: **mx**.

Looks like example.com has a couple of MX servers.

We look at the IP address for **mx1.example.com**.

If it matches the *client IP* we were given, we would return a **pass**! That's what would happen if the mail was legitimate.

But if this were a forgery attempt and the *client IP* wasn't found in either the **a** or the **mx** records, we would proceed to the last mechanism, **-all**, and return a **fail**.

If a temporary DNS error occurred, we'd return **error**.

Figure 2. SPF performs a simple DNS-based lookup for each incoming message.

nisms will return unknown, and your domain will be treated as though it did not have an SPF record at all.

Protecting Subdomains and MX Servers

Today, spammers forge domain names. Tomorrow, they might forge hostnames. They might try to joe-job your laptop by making up username@ibook.example.com. It's a good idea to protect your subdomains as well. You should start with your MX servers and move on to other hosts with A records. Here's why.

Bounce messages are sent with MAIL FROM: <>. The null sender address ensures that bounces don't themselves bounce and

BASIC SPF

A: the A mechanism means the IP address of example.com is permitted to send mail from example.com. If you want to say the IP address of some-other.com is permitted, you can say `a:some-other.com`. You can use as many A mechanisms as you want.

MX: the MX mechanism means the MX servers for example.com all are permitted to send mail from example.com. If you want to say the MX servers for some-other.com are permitted, you can say `mx:some-other.com`. You can use as many MX mechanisms as you want.

PTR: the PTR mechanism says if a host has a PTR record that ends in example.com, it is permitted to send mail from example.com. This would be a good choice for Yahoo, whose mail server names all end in yahoo.com. It would be a bad choice for a broadband provider like Comcast. If you want to say servers whose names end in some-other.com are permitted to send mail from example.com, you can say `ptr:some-other.com`. You can use as many PTR mechanisms as you want.

IP4: to say the class C network of 192.0.2.0 is permitted to send mail from example.com, you would write `ip4:192.0.2.0/24`.

ADVANCED SPF

Exists: the Exists mechanism takes an argument that expands to a domain name, and you can use macros. For example, `exists:%{ir}.%{l}._spf.example.com` might expand to `2.2.0.192.ceo._spf.example.com`. An SPF client would perform an A query on the expanded domain name, and if it got back any A record (for example, 127.0.0.2) Exists would result in a pass. You could use this technique to allow the special user ceo@example.com to send mail from a particular host, say 192.0.2.2, by creating an A record corresponding to the domain name above. Some people have written custom DNS servers to handle complex Exists queries. With Exists, the sky's the limit.

Include: if you send mail through another organization's servers, you should use the Include mechanism to point to their domain, so the SPF record is pulled in and expanded. For example, a vanity domain might use `include:isp.com` if it sends mail through ISP.com's mail servers. Any server permitted to send mail for ISP.com then is permitted to send mail for the vanity domain. You can include multiple other domains.

The modifiers Redirect and Exp: modifiers are different from the other mechanisms we've seen so far; they use equal signs instead of colons. Although mechanisms can repeat, you can have only one modifier per SPF record. Redirect is a modifier that works like Include, except the original query is replaced completely by the new query. Exp lets you define an explanation string. If an MTA rejects a forgery attempt, the explanation string appears in the SMTP error message that goes back to the original sender. You may have legitimate users who aren't using your SMTP servers, and SPF quickly can find out who they are. You also can set the explanation string to a URL that points to further information on how to configure mail clients correctly. All these mechanisms are described in detail at spf.pobox.com/mechanisms.html.

SPF AND CONVENTIONAL ANTISPAM METHODS

DNS blacklists or blocklists (DNSBLs): IPv4 space is 32 bits wide; 2^{32} is about 4.2 billion—4.2 billion grains of sand would just about fill a pickup truck. Imagine trying to paint each individual grain black or white. IP-based blacklists are a valiant effort, but they operate at too low a level. A good DNSBL has to decide whether an IP address is spammy and get it right for each of the 4.2 billion IP addresses. No wonder DNSBLs come and go—their maintainers burn out and give up.

Right-hand side blacklists: RHSBLs use domain names, whereas DNSBLs use IP addresses. Domain names are a much better way to identify entities on the Internet, but RHSBLs haven't been quite as popular as DNSBLs. Why not? Spam doesn't come from spammer.net. It's forged from yahoo.com. That's why SPF helps: if spammers send mail with their true names, blocking them becomes trivial.

Address verification: at MAIL phase, you can check the validity of the envelope sender by attempting to send a test message to it. If the test comes back user unknown, you might not want to accept the message. This is useful because spammers often make up addresses at random. But as address verification becomes more common, spammers can be expected to forge actual addresses—all the more reason to use SPF.

Signature solutions: PGP/GPG and S/MIME users sign their messages. Recipients can check signature validity by downloading keys from a key server. Other schemes have been proposed in which the DNS itself acts as the repository for public keys. These solutions are good because .forward files continue to work without modification. They are bad, however, because a message has to cross the pipe, costing bandwidth and CPU, before its legitimacy can be determined. In any case, a domain that uses these mechanisms still can use SPF to announce that any messages without a signature should be rejected.

Challenge/response: you don't want to send challenges to spam, especially not forged spam. If SPF tells you a sender address definitely was forged, you can junk the message without bothering to challenge it.

create a loop. When SPF sees the null sender address, it falls back to the hostname given in the HELO command. When your MTA sends a bounce message, it announces its hostname in the HELO command it sends. If that hostname has an SPF A mechanism listed, the message passes. So SPF prevents HELO forgery as well.

Traveling Mailman and the Forwarding Problem

SPF was designed to give the greatest benefit for the least cost. It tightens the rules in a way that makes it hard for bad people to do bad things, while not bothering the good people who do good things. Even so, some power users who have taken advantage of SMTP's lax rules may be inconvenienced by SPF. This section describes the two problems SPF causes power users and offers ways to work around them.

Most end users relay their outbound mail through their ISPs' SMTP servers. Most modern clients also support SASL authentication or POP-before-SMTP for users who need to phone home from outside the ISPs' networks. Users who always send mail through their ISPs' SMTP servers are automatically SPF-compliant and don't need to do a thing.

But some power users with an MTA on their laptop are used to originating mail from random IP addresses, bypassing their ISPs' SMTP servers entirely. SPF accommodates these users: the advanced mechanism (see the "Advanced SPF" sidebar) is a way to exempt certain users from being required to use their ISPs' SMTP servers. They can keep doing what they want.

Some power users have a dozen or more addresses that forward all over the place by using entries in /etc/aliases or .forward files. In classical forwarding, the envelope sender remains unchanged while the recipient address is rewritten. This becomes a problem, though, when the message arrives at the destination—it still has the original sender address, and SPF tests fail.

The workaround is easy, however; you simply need to switch to *remailing*, where the sender address changes as well. There are many ways to accomplish this. Read the SPF FAQ (spf.pobox.com/faq.html#forwarding) to pick the one that's right for you. Most end users have nothing to do with forwarding; only power users need to implement this workaround. If you have third-party service through an alumni, vanity domain or other commercial forwarding provider (such as pobox.com), you should expect them to implement remailing for you.

Stopping Spam: It's Part of the Solution

The primary goal of SPF is to stop forgery. I don't want to get any more spam from myself, and I certainly don't want you to receive any spam that claims to be from me. Worms and viruses tend to forge the envelope sender, too, and we can block them with SPF. And stopping forgery carries a bonus. When spammers are forced to use their true names, we can figure out which domains are legitimate and which are spammers. People already are doing this: a right-hand side block list (RHSBL) is the domain name version of a DNS block list (DNSBL). Spammers who aren't afraid of using their own domains end up on RHSBLs quickly, and they can be blocked that way. In an SPF world, RHSBLs will become more important and effective.

Why Do People Use SPF?

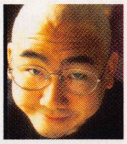
Big domains, including ISPs, banks and well-known brands care about controlling their trademarks. They have an obligation to protect their names. Altavista.com publishes an SPF record as do AOL and Oxford. More domains get on the bandwagon every day. Smaller domains publish SPFs simply because they don't want to be joe-jobbed.

On the receiving end, ISPs upgrade their MTAs and turn on SPF simply because it means less forgery—less spam, worms and viruses. Their bandwidth costs go down, too, because SPF lets them cut off the spammer before data is transmitted. They don't have to perform any cryptography or verify any signatures. SPF saves money.

Adoption

By the time this article is published, SPF support should be either bundled in or available as a downloadable plugin for the latest versions of SpamAssassin, Postfix, Sendmail, Exim and qmail. Commercial antispam vendors have committed to support SPF; Declude JunkMail, for one, reports that SPF is successfully blocking spam in the field.

If all goes well, the SPF standard will be published as an RFC in the near future. But thousands of domains, including some quite large ones, already publish SPF records. There's no reason to wait; you should publish SPF today.



Meng Weng Wong is founder and CTO of **pobox.com**, the e-mail-forwarding company, which celebrates its tenth anniversary this year. He is working on a science-fiction novel set on a planet where traditional fantasy magic turns out to be implemented, following Clarke's famous dictum, using nanotechnology.

Google Seeks Expert Computer Scientists

Google, the world leader in large-scale information retrieval, is looking for experienced software engineers with superb design and implementation skills and considerable depth and breadth in the areas of high-performance distributed systems, operating systems, data mining, information retrieval, machine learning, and/or related areas. If you have a proven track record based on cutting-edge research and/or large-scale systems development in these areas, we have plenty of challenging projects for you in Mountain View, Santa Monica and New York.

Are you excited about the idea of writing software to process a significant fraction of the world's information in order to make it easily accessible to a significant fraction of the world's population, using one of the world's largest Linux clusters? If so, see http://www.google.com/lj. EOE.

Google℠

Security Distribution for Linux Clusters

Here are the kernel mechanisms used in DSM to embed security information into IP messages in a transparent way. **BY IBRAHIM HADDAD AND MIROSLAW ZAKRZEWSKI**

This article is a follow-up to previous articles in *LJ* that discuss the Distributed Security Infrastructure (DSI) and the Linux Distributed Security Module (DSM) [see "Linux Distributed Security Module", *LJ*, October 2002, available at www.linuxjournal.com/article/6215, and "DSI: a New Architecture for Secure Carrier-Class Linux Clusters", available at www.linuxjournal.com/article/6053]. In this article, we focus on how we used IP options in DSM to send security information in a distributed environment for the process level of security. We discuss network buffer handling, adding hooks into the kernel, IP options and modifying IP headers. We then cover the network hooks in DSM and present some early performance results.

The DSI Project

The Open System Lab at Ericsson Research started the Open Source DSI Project to design and develop a cluster security infrastructure targeted at soft, real-time telecom applications running on Linux carrier-grade clusters. These clusters are expected to operate nonstop, regardless of any hardware or software errors. They must allow operators to upgrade hardware and software, kernel and applications, during normal operations, without any scheduled downtime and without affecting the offered services.

DSI originally was designed to offer carrier-grade characteristics, such as reliability, scalability, high availability and efficient performance. Furthermore, it supports several other important features, including a coherent framework, a process-level approach and support for both preemptive security and dynamic security policies.

One important feature of DSI is its process-level access control. Currently implemented security mechanisms are based on user privileges and do not support authentication checks for interactions between two processes belonging to the same user, even if the processes are created on remote processors. For telecom applications, only a few users run the same application for a long period of time without any interruption. Applying the above concept grants the same security privileges to all processes created on different nodes, which leads to no security checks for many actions through the distributed system. The granularity of the basic entity for the above security control is the user. For carrier-class applications, this granularity is not sufficient, and the need for a more fine-grained basic entity, the individual process, is required and thus supported in DSI.

The Distributed Security Module

The DSM is a core component of DSI that provides the implementation of mandatory access control within a Linux cluster. The DSM is responsible for enforcing access control and providing labeling for the IP messages with the security attributes of the sending process and node across the nodes of the cluster.

The DSM is implemented as a Linux module using the Linux Security Module (LSM) hooks. The development started using Linux kernel 2.4.17 along with the appropriate LSM kernel patch. The implementation was based on CIPSO and FIPS 188 standards, which specify the IP header modification.

One important aspect of the DSM implementation is its distributed nature. Access control in a cluster can be performed from a subject located on one node to a resource located on another node. Therefore, a need exists to transfer the security information between the nodes in the same cluster. The distributed nature of DSM provides location transparency of the security resources in the cluster from the security point of view.

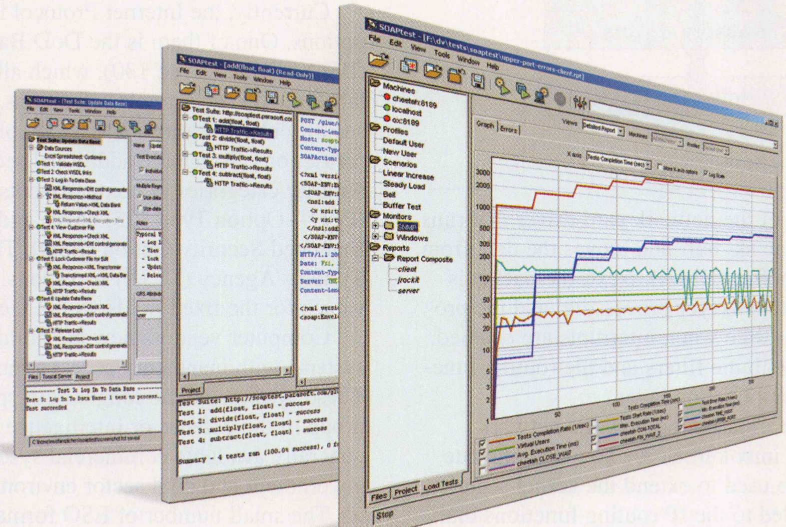
Network Buffer Handling

Here, we briefly discuss the topic of the network buffer handling to provide a better understanding of how security information is embedded into the network packet. We describe how the kernel handles network buffers starting from the application layer down to the hardware layer and vice versa.

Figure 1 shows the flow of the network packet in the kernel. Packet handling occurs in two cases, the incoming packet and outgoing packet. The outgoing network packet is handled as follows, starting from the application layer: the application prepares the data to be sent on the network; the application issues a system call to the kernel to send a packet; the packet, in the form of an sk_buff structure, goes through filters and routing functions inside the kernel; and the packet then is passed to the network driver that sends it to the network card (DMA).

The incoming network packet, starting from the network card, begins with the network card capturing the network packets either with its own address or the broadcast address; it then reads them to the network memory and generates an interrupt. The interrupt service routine, which is triggered by the hard-

It's automated. It's fast. And it's the most versatile web services testing tool ever.



SOAPtest
File Edit View Tools Window Help
Machines
cheetah:8189
localhost
ocx:8189
Profiles
Default User
New User
Scenarios
Linear Increase
Steady Load
Bell
Buffer Test
Monitors
SNMP
Windows
Reports
Report Composite
client
jrockit
server
Graph
Errors
Files
Project
Load Tests
Stop

Parasoft SOAPtest

The simple fact is, no other web service testing tool can do what SOAPtest can do. Or do it as fast and reliably. From ensuring functionality and interoperability to telling you how your web service is performing, SOAPtest automates all critical testing processes across the entire lifecycle of your web service.

But don't take our word for it... Go to www.Parasoft.com/SOAPtest and try it for free - no commitment, no obligation. If you like it (and we suspect you will) just let us know. We'll be more than happy to help you and your development team get up and running.

Server Functional Testing

- Automatically reads and tests client WSDL, to generate test cases.
- Records and captures HTTP traffic to build test cases.
- Flexible scripting with Java, JavaScript or Python
- Ability to emulate as server or client.
- Support for HTTP GET/POST.
- MIME attachments (per SOAP with Attachments specs)
- WS-Security, XML Encryption, XMLSignature Support.
- Data-driven testing through data sources.
- Works on new or already existing web services.

Load Testing

- Creates Windows and SNMP monitors to collect network information.
- Generates detailed reports for QA

Regression Testing

- Automatic Regression tests for functional tests suites and WSDL.

Conformance

- Conformance to WS-I Basic Profile 1.0
- WS-I Analyzer tool validates messages sent to and from a web service.

Protocol Support

HTTP 1.0
HTTP 1.1 w/Keep-Alive Connection
HTTPS
TCP/IP
JMS: supports testing of asynchronous Web services

Compatibility

- Works with any SOAP-compliant implementation, including Microsoft & IBM
- Windows 2000/XP
- Linux
- Solaris

Free Download...
Parasoft.com/SOAPtest

PARASOFT®
We make software work™

For Downloads go to www.Parasoft.com/SOAPtest. Or call 888-305-0041.

Copyright ©2003 Parasoft Corporation. All rights reserved. All Parasoft product names are trademarks or registered trademarks of Parasoft Corporation in the United States and other countries. All other marks are the property of their respective owners.

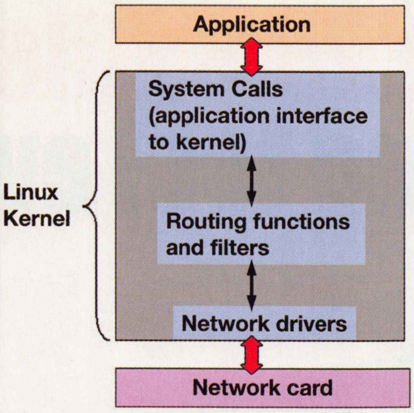

Application
System Calls (application interface to kernel)
Linux Kernel
Routing functions and filters
Network drivers
Network card

Figure 1. Network Packet Flow in the Kernel

ware interrupt and is a part of the network card driver that runs inside the kernel, allocates an sk_buff and moves the data from the card memory into this buffer (DMA). Next, the packet is put on the CPU queue for upper-layer processing, and the processing is deferred to a later time when interrupts are enabled. Finally, the packets go through the filters and the routing functions and are passed to the application layer.

Based on the generic information on how the network buffers are handled in the Linux kernel, we now demonstrate how this information can be used to extend the kernel security. We look into the hooks added to the IP routing functions that allow us to manipulate the IP packets and add extra security to the IP messages.

Adding Network Security Hooks

The security module can influence the routing decision based on the security hook implementation. A few things need to be remembered when programming the routing hooks, because those hooks are executed as normal kernel functions for every packet coming in and out the kernel.

A module that registers a function must specify the priority of the function within the hook. The net filter hooks are called from the kernel code in the order of priorities. The user functions are free to manipulate the IP packet. The user function must return one of the following values in order for the networking code to decide what to do with the packet:

1. NF_ACCEPT: do nothing and let the packet go through the network stack.
2. NF_DROP: drop the packet. The packet is not passed for further processing.
3. NF_STOLEN: the packet has been taken. The packet is not passed for further processing.
4. NF_QUEUE: queue the packet for user-space handling.
5. NF_REPEAT: call this hook again.

This function shows us how the packet is manipulated before it enters the system and before it is sent out. What we are still missing is: what kind of information or options can we add to the packet? How? And, will the changes coexist with the current implementation? We answer these questions in the following sections.

IP Options

A little-known fact about Internet Protocol is that an IP packet can contain a variable amount of extra information (maximum of 40 bytes) following the standard 20-byte header. These extension bytes are called IP options, and some of the options are defined to carry security information.

Currently, the Internet Protocol includes two security options. One of them is the DoD Basic Security Option (BSO—Option Type 130), which allows IP datagrams to be labeled with security classifications. This option provides 16 security classifications and a variable number of handling restrictions. To handle additional security information, such as security categories or compartments, a second security option (ESO—Option Type 133) exists and is referred to as the DoD Extended Security Option (ESO). The Defense Information Systems Agency (DISA) is responsible for administrating the values for the fixed fields within these two options.

Computer vendors now are building commercial operating systems with mandatory access controls and multilevel security. These systems are no longer built specifically for a particular group in the defense or intelligence communities. They are generally available commercial systems for use in a variety of government and civil sector environments.

The small number of ESO format codes cannot support all the possible applications of a commercial security option. The BSO and ESO were designed to support only the United States DoD. Commercial IP Security Option (CIPSO) has been designed to support multiple security policies. The Internet draft provides the format and procedures required to support a mandatory access control (MAC) security policy.

The IP options used to label packets in our implementation are based on the FIPS 188 standard and the Commercial IP Security Option (CIPSO) draft. In our implementation, the IP header is changed using these standards, so we can add the security information to the IP header and send it over the network.

IP Options in DSM

The security information we want to transfer using IP options are Security ID (SID) and Security Node ID (NID). The DSM modifies every IP packet by supplying our security information as its IP options. Figure 2 shows the format of the modified IP header.

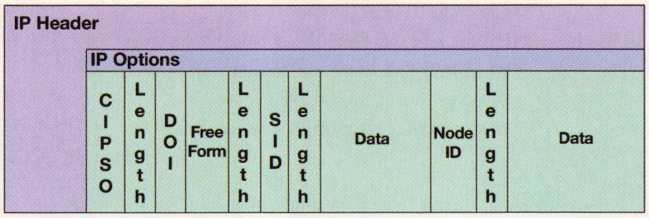
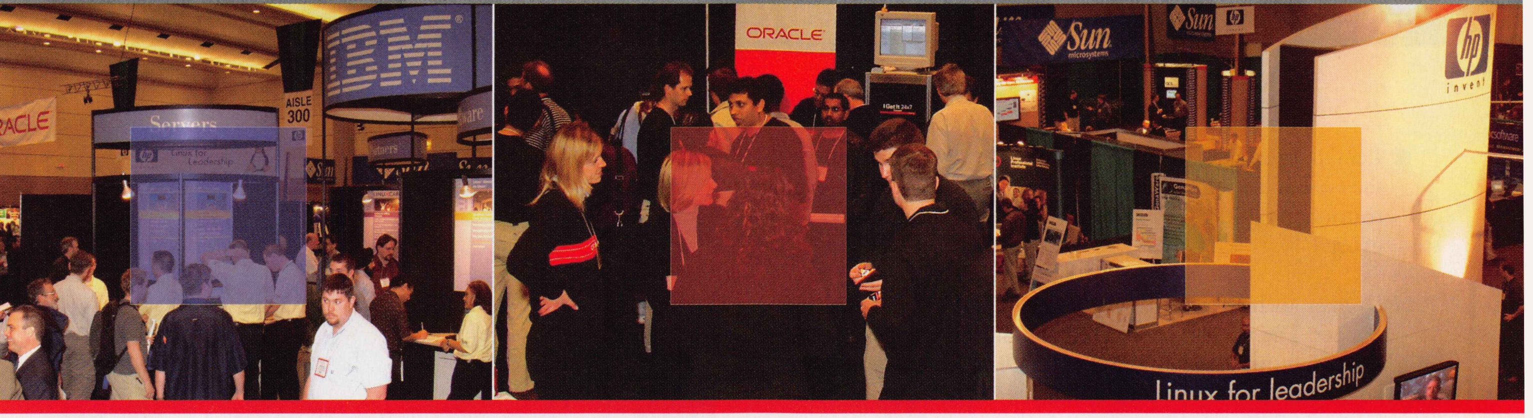


Figure 2. Security Options in the IP Header

Here is a list of header options:

- CIPSO: one octet, with a value of 134.

Your Connection To Today's Applications and Solutions



REAL WORLD LINUX 2004
CONFERENCE & EXPO

April 13-15, 2004

Metro Toronto Convention Centre North Building Toronto, Canada

A Key Conference And Trade Show Exclusively Dedicated To Linux And Open Source

EVERYTHING YOU WANT TO KNOW ABOUT LINUX AND OPEN SOURCE UNDER ONE ROOF

Learn Valuable intelligence on key issues at a dynamic conference
Interact One-on-one with leading vendors
Navigate RWL04 is your compass to success with Linux
Ultimate Arena for users, advocates and those that want to learn about Linux
Xceptional Opportunity to learn about LINUX NOW!

Platinum Sponsors
Novell IBM hp invent

Silver Sponsor
redhat

Bronze Sponsors
IMP solutions TOPSPIN ILOG Changing the rules of business Candle Managing what matters most @lliance Technologies Avocent EQUINOX an Avocent Company RLX technologies smarter computing. intelligent management

Premium Media Sponsor
COMPUTERWORLD NetworkWorld

Association Platinum Sponsor
CATAAlliance

Conference Sponsor
IDC Analyze the Future

Certification Sponsor
Linux Professional Institute

Media Sponsors
SVN silicon valley north MARCOM OSDN LinuxWorld Magazine geek street.ca everything geek across Canada LINUX.COM LINUX MAGAZINE LINUX JOURNAL

Supporting Associations
YTA APCC COMMUNITECH CANARIE Ontario Society of Professional Engineers CAIP ACFI Novell Users International Toronto Chapter Society of Internet Professionals Wired Woman Society Linux.ca CLUE: the Canadian Linux users' exchange CIPS Toronto Section IEEE Canada Toronto Section LINUXCHIX

TO REGISTER OR EXHIBIT GO TO www.realworldlinux.com
TOLL FREE: 888.823.7586 - info@realworldlinux.com - exhibit@realworldlinux.com

Produced by Real World Linux Inc.

when registering quote source code LJ2

- Length: one octet, the total length of the option including the type and length fields. With the current IP header length restriction of 40 octets, the value of this field must not exceed 40.

- Domain of Interpretation (DOI) Identifier: unsigned 32-bit integer. The value 0 is reserved and must not appear as the DOI identifier in any CIPSO option. Implementations should assume the DOI identifier field is not aligned on any particular byte boundary.

- The CIPSO Domain of Interpretation (DOI) Field, or the Security Tag Set Name under FIPS 188: set to hexadecimal 10001000. This DOI value was selected arbitrarily as there currently is no relevant regulatory activity in this area.

- Free Form: one octet, indicates that the following fields are new fields undefined in standard (therefore free). The value is 7.

- Length: one octet, indicates the total length of all tags.

- Tags (SID, NID): CIPSO uses sets of tags to contain the security information relevant to the data in the IP packet. Each tag begins with a tag type identifier followed by the length of the tag; it ends with the actual security information to be passed.

- SID tag: tag id: one octet (value 3), tag length: one octet (value 6), tag data: 32-bit value of sid.

- NID tag: tag id: one octet (value 6), tag length: one octet (value 6), tag data: 32-bit value of nid.

- The IP option we use is CIPSO. Those fields are not defined by the standard, so they can be used in the way we define.

- The Domain of Interpretation (DOI) and the Free Form (FIPS 188 standard) mean that the following fields are new fields undefined in the standard, therefore they are free.

DSM Network Hooks

We used the LSM security hooks in the DSM to add our security labels to the IP messages. We now demonstrate how we achieved this by presenting an example of an application that sends a packet over the network by writing to a socket. The application uses some of the library calls. At one point, a system call is generated that passes the message to the Linux kernel. The entry point to the kernel socket implementation is the function sys_socketcall(), located in net/socket.c. In the chain of calls, the sock_sendmsg() function (Listing 1) in net/socket.c is executed.

One of the first actions in the function is to execute the security hook (`security_ops->socket_ops->sendmsg(...)`). This hook ends up in the DSM socket hook that modifies the IP packet, as shown in Listing 2.

The function dsi_options_fill sets up the security information to the buffer as specified in the previous paragraph. Later, in subsequent functions, this security information is attached to the IP message as options. The SID is derived from the socket security ID, and the NID is global for the whole node—there is no need to pass it as a parameter to the function.

Listing 1. sock_sendmsg()

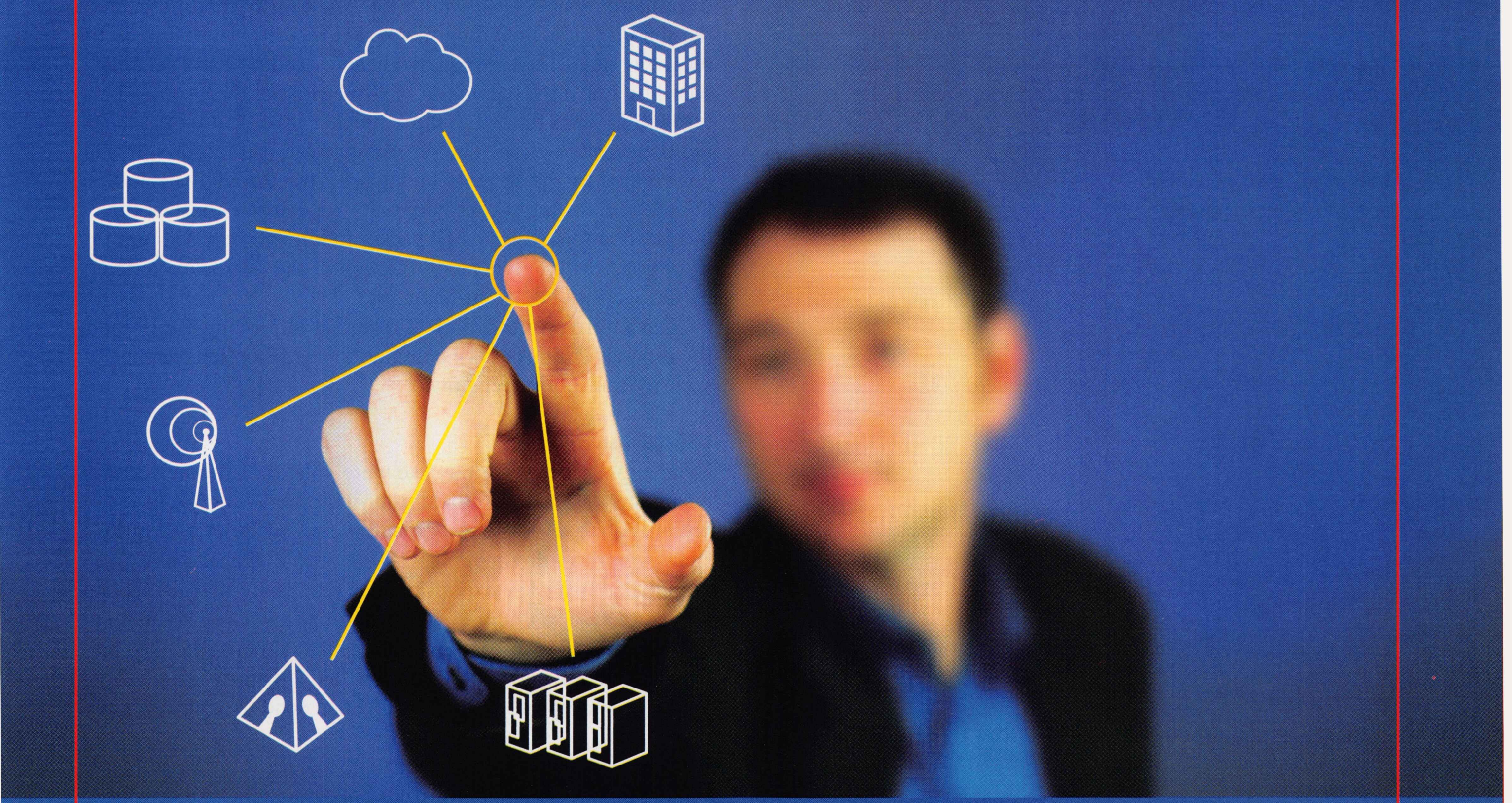
```
sock_sendmsg
(struck socket *sock, struct msghdr *msg, int size)
{
  int err;
  struct scm_cookie scm;

  err =
    security_ops->socket_ops->sendmsg(sock,
                                   msg, size);
  if(err)
    return(err);
  ...

}
```

Listing 2. dsi_socket_sendmsg()

```
int dsi_socket_sendmsg(struct socket *sock,
                     struct msghdr *msg, int size)
{
...

inode_security_t *isec;
struct sock sk;
struct ip_options *opt =  NULL;
int optlen = NSID_BASE_LEN + NSID_SSID_LEN +
             NSID_NODEID_LEN; //8 +_6 + 6
unsigned char optptr[optlen];

  ...

  sk = sock->sk;
  opt = sk->protinfo.af_inet.opt;
  dsi_options_fill (isec, optptr, optlen);
  dsi_ip_options_get(&opt, optptr, optlen);
  opt = xchg(&sk->protinfo.af_inet.opt, opt);

...
}
```

After this action, the modified packet with the security information added is forwarded for normal processing in the kernel and finally is sent over the network. At the receiving side, the incoming messages are stored in the sk_buff structures and preprocessed in a series of functions and hooks. One of these functions is ip_options_compile (Listing 3) in /net/ipv4/ip_options.c, where the options are processed. For the CIPSO case, the security hook decode_options is called. This hook is replaced by the DSM dsi_decode_options hook, where the security parameters (SID, NID) from the incoming packet are read and stored in the security structure attached to this sk_buff. The sk_buff buffers, populated with the security information, are attached to the receiving socket queue, where they are waiting to be read by the receiving application. In order to read them, the application issues the

ALL POINTS COVERED.



» SECURITY » WIRELESS » COLLABORATION & VoIP
» INFRASTRUCTURE & SERVICES » PERFORMANCE & DATA CENTER

Your role is expanding.

As business infrastructure moves to the forefront, your company relies on you to keep existing systems at peak performance while tackling a broad range of new requirements. Securing wireless networks, implementing collaboration technologies, improving regulatory compliance, ensuring business continuity—all while doing more with less.

Enterprise challenges demand end-to-end solutions.

As networking and communications professionals, you need to focus on an end-to-end, system-wide approach when building and optimizing your business infrastructure:

- **» Security**
- **» Wireless**
- **» Collaboration and VoIP**
- **» Performance**
- **» Data Center and Storage**
- **» Infrastructure and Services**

Visit INTEROP.com for details and registration.
Save up to $200—Register by April 9.
Use Priority Code: ADBKZ1ND

Put it all together.

Only NetWorld+Interop brings you the latest strategies, techniques and products for every point in your infrastructure—and shows you how they can add up to an integrated, end-to-end solution that meets every requirement on your list.

Make the connection at NetWorld+Interop.

For 17 years, NetWorld+Interop has helped networking professionals take their enterprises and their careers to the next level. At NetWorld+Interop Las Vegas 2004, you'll make the industry's best ideas and latest technologies relevant to your needs, and discover that even the toughest networking challenge is all in a day's work.

NETWORLD℠
+INTEROP®
LAS VEGAS • MAY 9-14, 2004
EXHIBITION: MAY 11-13, 2004

MediaLive™ INTERNATIONAL

Copyright © 2004 MediaLive International, Inc., 795 Folsom Street, 6th Floor, San Francisco, CA 94107. All Rights Reserved. MediaLive International, Networld, Interop, and associated design marks and logos are trademarks or service marks owned or used under license by MediaLive International, Inc., and may be registered in the United States and other countries. Other names mentioned may be trademarks or service marks of their respective owners.

The End-to-End Networking and Communications Event

Listing 3. ip_options_compile ()

```
int
ip_options_compile (struct ip_options *opt,
                    struct sk_buff *skb)
{
unsigned char *pp_ptr;
unsigned char *optptr;

  ...

  case IPOPT_CIPSO:

    if(security_ops->ip_ops->decode_options(skb,
                                    optptr, &pp_ptr)
      goto error;
      break;
  ...
}
```

Listing 4. sock_queue_rcv_skb ()

```
int
sock_queue_rcv_skb (struct sock *sk,
                    struct sk_buff *skb)
{
int err=0;

  ...

  err=security_ops->socket_ops->sock_rcv_skb (sk,
                                              skb);

  if(err)
    return (err);
  ...
}
```

system call sys_socketcall (), as it did for the sending packet. The call once again goes through the DSM security hook, where the receiving socket security ID is validated against the sk_buff security of the incoming packet. If the socket is not allowed to receive the packets with a given security ID, then those packets are dropped. Listing 4 shows the kernel function in include/net/sock.h.

As we can see, the security hook sock_rcv_skb is called. This hook then is replaced by the DSM function dsi_sock_rcv_skb when the DSM is loaded. In this function, the security validation is performed. From the example code we can see work needs to be done to manipulate the security labels.

Performance Measurements

We performed several benchmarking tests in order to verify whether adding options to the IP header affects the overall performance and by how much. One test was to send a UDP packet between nodes of the cluster and measure the performance degradation that consists of the packet security modification on the sending side, including the packet security extracting on the receiving side. The average overhead of adding extra security based on our implementation is 30%. Most of the overhead (around 25%) is related to the IP packet modification based on the IP security option. The remaining overhead (around 5%) is contributed by the security hooks infrastructure in the Linux kernel, such as the socket hooks. As we can see, most of the overhead is related to the IP packet modification based on the IP options, with only a small fraction of the overhead caused by the security hooks infrastructure.

Our future efforts will be directed at improving the IP modification algorithms as we continue to use IP options as the security transport mechanism.

Conclusion

By changing the IP options, we were able to distribute security information to nodes of the cluster with the DSM. We have optimized the IP packet modification and our primary results show significant improvements—the 30% overhead has dropped to 14%. These performance results are promising, and we see more opportunities for further optimizations to attain a lower overhead. Nevertheless, the results demonstrate the challenges facing the development of efficient distributed security. We hope you try out DSI and DSM and send us your feedback.

Acknowledgement

David Gordon, co-op intern from Sherbrooke University, for his contributions to DSM.

Ibrahim Haddad, contributing editor of *LJ*, is a researcher in the Research & Innovation Unit at Ericsson Research in Montréal, Canada. He contributed to two of Richard Peterson's books, *Red Hat Linux Pocket Administrator* and *Red Hat Enterprise Linux & Fedora Edition: The Complete Reverence* (DVD edition), published by McGraw-Hill/Osborne.

Miroslaw Zakrzewski works for Ericsson Canada in Montréal, developing the new-generation CDMA systems. He can be reached at Miroslaw.Zakrzewski@Ericsson.ca

REFERENCES

DSI and DSM Home Page: **www.linux.ericsson.ca/dsi**

FIPS 188: **csrc.nist.gov/publications/fips/fips188.html**

Linux Packet Filter: **www.linuxjournal.com/article/4852** and **www.linuxjournal.com/article/5617**

LSM: **lsm.immunix.org**

Network Buffers: **www.linuxjournal.com/article/1312**

Open System Lab: **www.linux.ericsson.ca**

SE Linux: **www.nsa.gov/selinux**

Power to the People

Open-source philosophy, not to mention technology, is infiltrating more and more aspects of our daily lives. **BY HEATHER MEAD**

Back in November 2003, Doc Searls posted a short piece on the *LJ* site (www.linuxjournal.com/article/7239) that outlined which presidential campaign Web sites were using open-source components. This seems particularly relevant given the significance of the Internet to the 2004 presidential race. As Howard Dean's early campaign demonstrated, making the Internet key to an organization can turn a lot of separate grassroots initiatives into a large and powerful networked movement.

By the time you read this month's On the Web, we'll know who the 2004 Democratic presidential candidate is. Although this is a big story, Doc continues to be interested in the story behind the story—how the campaigns are adopting and adapting open-source philosophy and technology. He describes his mission as learning:

> what IT workers in the pressure-cooker conditions of political campaigns might teach IT professionals everywhere about the resourceful use of Linux, free software and open-source development methods. What works best? What doesn't work at all? How do you develop and apply solutions to problems all over the country with widely varying participants and circumstances?

Prior to LinuxWorld New York in January 2004, Doc traveled to the Vermont headquarters of the Dean campaign. As he writes in "Lessons from the Campaign Pressure Cooker" (www.linuxjournal.com/article/7372), he encountered both open-source software and what former Dean Campaign Manager Joe Trippi called "open-source politics". Joe used to work for Ian Murdock, cofounder of Debian. Essentially, the Dean campaign used its extensive grassroots support to create its own networked market. How far this market takes Dean is unknown right now, but its success in 2003 seems to indicate that a new phase of electoral politics has begun.

Perhaps this shift in the campaign landscape is connected to the shift taking place in the larger realm of supply and demand. In "The New Economy Hack: Turning Consumers into Producers" (www.linuxjournal.com/article/7345), Doc discusses how "consumerism is a red herring....It isn't about what you and I invent and contribute to the marketplace. It's about what Sony and Panasonic and Nikon and Canon produce and distribute through retailers for us, the mass market, to consume constantly." Some new computer technologies and the overall continuing drop in cost of computer equipment, however, are allowing users to have more control over what they create and use in their daily lives. Doc traces all this back to the Linux economy hack, "because Linux is something that happened when demand started to supply itself".

Whether looking at political campaigns and the importance of the Internet, media coverage and the rise of the blog or consumer electronics and the increasing availability of software that lets users make their own music, it's clear that the do-it-yourself freedom at the heart of open source is spreading. If you'd like to keep up with Doc's findings and musings, subscribe to his biweekly SuitWatch newsletter at the *Linux Journal* home page.

Heather Mead is senior editor of *Linux Journal*.

EQUALIZER
EXT. INT.
SERIAL

Win the battle with server costs.

(Load balancing: the affordable way to attack server capacity limits.)

If traffic is jamming your servers, the answer isn't more servers. It's more traffic capacity. Get it with the Equalizer, the affordable load balancer that distributes requests between servers here, there or anywhere on your network. So end the assault on your IT budget. Call **650-969-6000** today or visit **www.coyotepoint.com**

COYOTE POINT SYSTEMS INC

©2002 Coyote Point Systems, Inc.

AOL for Linux?

I have a dual-boot (GRUB) setup with Red Hat Linux 9.0 on one partition and Windows XP Professional on another. I also am running AOL 9.0 broadband software for my ISP. Is there a driver for the AOL Broadband Blaster modem that I can use to run Red Hat 9.0 with Internet services through my AOL broadband connection? I want to avoid changing Internet services and buying a new DSL modem, but I can't get Red Hat to see the Broadxent modem, and Linux doesn't recognize my AOL broadband software.

--
Natosha, NZimardo@aol.com

You cannot run AOL's software under Linux, so you will not be able to access AOL services through traditional means. However, for general Internet access, many Broadxent DSL modems have both USB and Ethernet interfaces. If you can switch to the Ethernet interface and install an Ethernet card in your PC, you should be able to access the Internet from both operating systems. If your DSL modem does not have an Ethernet interface, you need to replace it with either a USB device that Linux supports (such as from Alcatel or ECI) or a device with an Ethernet interface. If you cannot get AOL to agree to allow you to do this, you should consider an alternate service. Linux desktop deployments are becoming more common today, especially with the efforts of Lindows and other distribution vendors working with retailers to package their products into inexpensive solutions.

--
Chad Robinson, crobinson@rfgonline.com

Collecting the Ultimate Linux Box

I have followed Glenn Stone's "Ultimate Linux Box" series on the Web and in print and decided to build my own version. I have found the economical way is to build it piece by piece; I buy a piece every month and put it away until I get all the pieces, then I'll put it together. After doing a lot of research I came to the conclusion that 1) modular is the way to go, and 2) 64-bit is the future. I have thus elected to build an AMD Opteron single CPU system running Mandrake Linux, 9.2/64. I am thinking of buying the ASUS SK8N Motherboard (2GB of RAM, twin Maxtor 120GB SATA HD, 3.5 floppy, 5.25 floppy (legacy work), CD-RW (48x12x48x) and, with luck, a DVD player. Is this a problem board on Linux? If it really is a problem board, can you suggest a solid alternative board I could buy?

--
S.W. Bobcat, swbobcat@hotmail.com

If you have problems with this motherboard, it is more likely to be a driver support issue than anything else. Most board-related issues can be resolved with simple workarounds, such as the noacpi boot option. This motherboard in particular was used in several SPEC.org benchmark tests in 2003, so I doubt there is an issue that cannot be resolved in some fashion.

However, I would advise you to consider your 64-bit choice carefully. Unless you are doing some serious rendering or mathematical computation, the 64-bit platform is unlikely to provide you significant benefit. In fact, you may find that some applications run slightly more slowly; many benchmarks today are showing that unless a workload is optimized for a 64-bit platform, it does not perform to its full ability. 64-bit computing ideally is suited for a variety of computational workloads or for things that require wider integers for indexing facilities, such as databases. Descent 3 *won't run any faster.*

There is no question that many environments can benefit from this change, as long as you recompile everything to support it. Is yours going to be one of these? On the other hand, the Opteron itself is a great processor, and if you are making the choice based on HyperTransport, for instance, then kudos!

--
Chad Robinson, crobinson@rfgonline.com

I'm afraid, I can't agree with your modular assertion. PCs have become so complex that you take the risk of ending up with incompatible chipsets or a system with subtle bugs. This doesn't mean you will, but as PCs have become more than a thousand times faster, the mix-and-match approach is a lot more likely to hit some subtle timing incompatibility between two components. As for 64-bit being the future, yes, although you are not likely to need 64 bits for what you are going to do. A wait-and-see approach for 64 bits wouldn't be unreasonable in your case (and in the meantime, you could buy a 32-bit system). In addition, you will not end up with a working system before you buy and assemble the last piece, and even then, it's only if you are lucky and everything does piece together. You will not only pay more for all the components bought separately, but the price of the first components you bought will have dropped by the time you buy the last one. I really recommend that you buy a prebuilt system unless you are really looking to do hardware tinkering and are ready to swap components and meet potential incompatibilities.

--
Marc Merlin, marc_bts@google.com

It's better to save up for the parts and buy them all at once. That way the warranty for the first part isn't ticking away while you collect the rest. And, as Marc points out, PC hardware does tend to get cheaper over time. If you're building your own Linux system, explore the Web sites of Linux system vendors and see what hardware they use. It's likely to be stable and compatible.

--
Don Marti, dmarti@ssc.com

How to Use #include

I am trying to access ports using check_region() . I have included the file /linux/resource.c to do this. But when I try to compile the program, I get the following error:

```
In file included /usr/include/linux/sched.h
                                  /linux/resource.c
/usr/include/linux/timex.h :field "time" has incomplete
type
```

where time is of type struct timeval declared in the file timex.h.

Kindly tell me the possible reason and its solution.

--
Ashutosh Sharma, catchwavesin@yahoo.com

resource.c is a source code file that contains all of the function implementations, including the one you are trying to use. I believe you should be including <linux/ioport.h> instead, which contains the definitions for these functions, not the implementations.

--
Chad Robinson, crobinson@rfgonline.com

Big Pictures from The GIMP?

I use The GIMP for almost everything in picture editing, and I have created some large pictures, say 24 × 24 inches. I basically use them as first-draft examples before finally painting my subject on canvas. I have an HP 5650 printer that obviously cannot print such large paper sheets. So the results are that only the upper-left corner of the picture is printed out. What I'm looking for is the ability to print the entire picture in its correct size but spread out on several 8 × 10 sheets of paper that I then can paste together to create the correctly sized pictures. Does any such software exist in the Linux world, or must I purchase HP's Plotter to achieve this result at a greater cost?

--

Paul Godin, linuxstuff@istop.com

Perhaps I'm confused, but your problem sounds as though it has a simple solution. You simply need to crop successive sections of the image so you can print one tile at a time. You can do this manually with The GIMP. You also can create a script to do this automatically, using the ImageMagick toolkit. There is a tool called convert(1) in this suite that not only allows you to convert images to a format more suitable for printing (such as PostScript or PCL) but also to crop portions from the image. See the man page for this application for details, but I believe you probably want to use the -crop option.

--

Chad Robinson, crobinson@rfgonline.com

You can print your file in PostScript and then use this program to split it onto multiple pages: www.ctan.org/tex-archive/support/poster. *It also prints the marks you need to cut the pages and paste them together.*

--

Marc Merlin, marc_bts@google.com

What you are trying to do is called a mosaic. There are many tools to do mosaics with The GIMP; take a look at registry.gimp.org. *Also, bear in mind that your printer may have blind printing areas where it is not possible to print, perhaps in margin areas, so take this into account when doing the mosaic. You also may need to do some traditional cutting on each sheet of paper.*

--

Felipe Barousse Boué, fbarousse@piensa.com

Making Red Hat 9 into a Firewall

I am trying to set up a box as a firewall so routing is enabled. I have two Ethernet cards for this box. I am running FWbuilder 1.1.1 on Red Hat 9. My question concerns the kernel route tables. The situation is the external address is on eth0 192.168.1.2, and the internal address is on eth1 172.10.10.252. My default gateway on eth0 should be 192.168.1.1; however, the kernel makes the default 192.168.1.2 and will not release this route. This now has routed to the 192,168.1.0 side, but it stops at eth0 and cannot find a way to 192.168.1.1. The internal side works great, though. Besides rebuilding the kernel, how can I set

Stop Hunting for a Service Provider...
Host Your Site with Hurricane Electric!

Flat Rate Gigabit Ethernet
1,000 Mbps of IP
$27,000/month

Full 100 Mbps Port
With Full Cabinet
$2,500/month

Colocation Full Cabinet
Holds up to 42 1U servers
$400/month

Order Today!
email sales@he.net or call 510.580.4190

HURRICANE ELECTRIC
INTERNET SERVICES

the default gateway to be 192.168.1.1?

--

Joe Golden, jgolden3@csc.com

Edit the /etc/sysconfig/network-scripts/ifcfg-eth0 file and add a line like `GATEWAY='192.168.1.1'`*, and then restart your network with the service network stop/service network start commands or reboot your machine. This fixes the default gateway route of the eth0 card to 192.168.1.1.*

--

Felipe Barousse Boué, fbarousse@piensa.com

Switching between Running X Sessions

How can I have multiple X sessions running at the same time? That is to say, if I'm logged in as root and also log in on Ctrl-Alt-F2 and give the command `startx -- :1` to fire up X, everything works fine. If I go back to Ctrl-Alt-F7, all is fine there too. But, when I go to Ctrl-Alt-F2 again, X has crashed there, but it still is up on F7. Is there some command that I can give so X stays up on F2 when I'm going back and forth?

--

Bjarni Valsson, bjarniv@hotmail.com

X does not run on F2 (pty 2). Each X instance you start creates its own pty, hence the switch to pty 7. If you go back to pty 2, you actually can put X in the background (Ctrl-Z, then type `bg`*) and continue using that console for other tasks. If you start up another copy of X, and X already is running, it creates a new pty (now 8, in this case). You must then press Ctrl-Alt-F8 to switch to that copy of X. Remember, X is a user-space application. Two running X processes should not interfere with one another unless you are doing something odd with your hardware; certain video card driver settings might cause trouble.*

--

Chad Robinson, crobinson@rfgonline.com

DHCP under Knoppix but Not Debian

I recently installed a Linksys WPC11 wireless card on an older Gateway 433 Celeron machine running Debian Woody 3.0. The card works fine only after I log in as root and type `pump -i wlan0`. I followed the instructions to the letter and even modified the specified files to read DHCP as the instructions stated. The interesting part of this equation is that I ran Knoppix on the same machine out of curiosity, and Knoppix correctly identifies the card, and it automatically gets an IP address. Any idea what I can do to make this card pull an IP automatically?

--

Wes Reneau, wes@rose.net

Woody 3.0 is an exceedingly old version of Debian (mid-2002), as far as Linux distributions go. You should consider upgrading to the latest distribution. Knoppix is based on Debian, but you are undoubtedly using a fairly recent version. This is probably nothing more than an issue with one of your boot scripts, but without viewing them it's difficult to tell.

--

Chad Robinson, crobinson@rfgonline.com

"dump"ing an ext3 Filesystem?

I regularly back up all my machines using amanda and dump. The backup is to a dds-3 tape drive on a Red Hat 7.2 machine (servred72). I upgraded a Debian 3.0 machine to sarge and had some problems. I wanted to amrestore on the Debian 3.0 machine, but amrestore requires root privilege to run. I was prevented from the restore because of security mechanisms. The appropriate amandahosts files were modified to include root. Root access was prevented in the LAN. Two questions: 1) dump works on only ext2 filesystems. Is there a newer version to support the journaling filesystems? 2) What could I do to relax security temporarily and allow amresotre to work as root across the LAN?

--

Alan Polinsky, polinsky@acm.org

A dump that's recent enough should work with ext3, because ext3 is ext2 with a new feature flag on the filesystem and a special hidden inode with the journal, which you can but don't need to backup. You do need to quiesce the journal when you take a snapshot for backup, but dump has support for that, as hinted by the sf.net Project page, **sourceforge.net/projects/dump**, *which says "Dump/Restore ext2/ext3 filesystem backup".*

--

Marc Merlin, marc_bts@google.com

Many on-line help resources are available on the *Linux Journal* Web pages. Sunsite mirror sites, FAQs and HOWTOs can all be found at **www.linuxjournal.com**.

Answers published in Best of Technical Support are provided by a team of Linux experts. If you would like to submit a question for consideration for use in this column, please fill out the Web form at **www.linuxjournal.com/lj-issues/techsup.html** or send e-mail with the subject line "BTS" to bts@ssc.com.

Please be sure to include your distribution, kernel version, any details that seem relevant and a full description of the problem.

Linux
2.6 Kernel
Linux
Real Time
24.1.10
Single Board Computer
With Analog & Digital I/O

PCM - 53E52

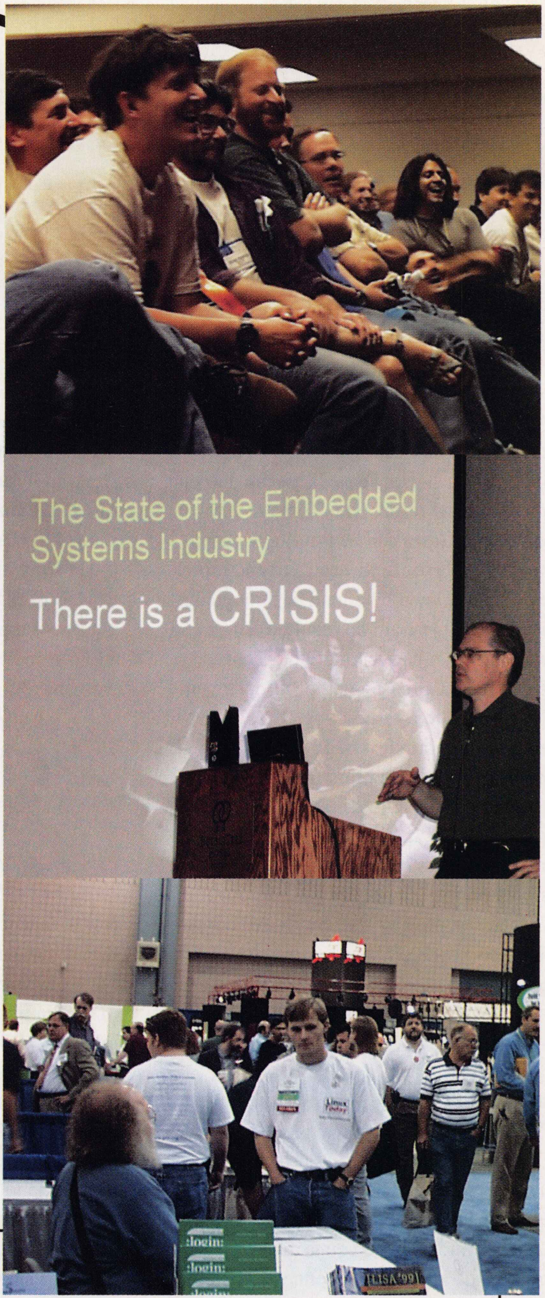
- 100% PC Compatible
- Compact Size (4.0" x 5.75")
- Low Power, No Fan Required
- AMD SC520 133MHz CPU (P120)
- 10/100Base-T Ethernet
- Supports DOC Flash Disk
- PC/104+ 32-bit Bus Expansion
- Onboard SDRAM from 16 to 128 MB
- Keyboard & Mouse Device Interface
- Four Serial Ports & One Parallel Port
- Two IDE Device & Two FDD Device Interface
- Optional Enclosure & Power Supply
- Analog & Digital I/O Daughter Board Option

Starting at
$320.00 USD

Since 1985
OVER
19
YEARS OF
SINGLE BOARD
SOLUTIONS

EMAC, inc.
EQUIPMENT MONITOR AND CONTROL

Phone: (618) 529-4525 • Fax: (618) 457-0110 • Web: www.emacinc.com


The State of the Embedded Systems Industry
There is a CRISIS!
;login:
;login:
LISA '99

USENIX
THE ADVANCED COMPUTING SYSTEMS ASSOCIATION

The USENIX Association brings together developers, engineers, system administrators, researchers, and technicians working on the cutting edge of computing systems. Our conferences provide a forum for the presentation and discussion of the most advanced development and implementation information on technical topics ranging from secure computing and system design to file and storage technology.

Membership offers many benefits, including:
- Subscription to the highly respected magazine, *;login:*
- Access to all Conference Proceedings online
- Discounts on USENIX conferences and materials
- Discounts on books and publications such as *SysAdmin, Linux Journal*, O'Reilly and No Starch Press books, and more.

For more information on membership benefits and conferences, visit **http://www.usenix.org**.

USENIX EVENTS 2004-2005

VM '04
Third Virtual Machine Research and Technology Symposium
May 6–7, 2004
San Jose, CA

USENIX '04
USENIX Annual Technical Conference
June 27–July 2, 2004
Boston, MA

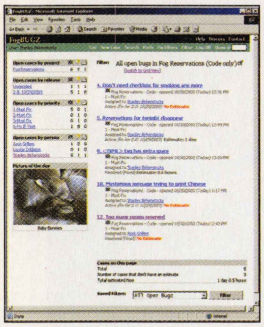
13th USENIX Security Symposium
August 9–13, 2004
San Diego, CA

LISA '04
18th Large Installation Systems Administration Conference
November 14–19, 2004
Atlanta, GA

OSDI '04
Sixth Symposium on Operating Systems Design and Implementation
December 6–8, 2004
San Francisco, CA

USENIX '05
USENIX Annual Technical Conference
April 10–15, 2005
Anaheim, CA

FogBUGZ 3.0 for UNIX



FogBUGZ 3.0 is a Web-based software project management system. Created to be a database of cases, which can be feature requests, traditional bug reports or customer e-mails, every case is assigned to one person who must resolve it or forward it to someone else. Cases can be prioritized, documented, edited, estimated, searched and more. Cases can be entered by e-mail or through the Web interface; there are no required fields, and anyone can edit a bug report. Screenshots, sample files and almost any type of document can be attached to a case in FogBUGZ, and Unicode is supported so bugs can be entered in any language. In addition, FogBUGZ can be integrated with source code managers, such as VSS, CVS, Vault and Perforce, and bidirectional links between check-ins and bugs may be maintained. FogBUGZ 3.0 for UNIX can run on Red Hat, SuSE, Mandrake, Debian and FreeBSD.

CONTACT Fog Creek Software, 535 Eighth Avenue, 18th Floor, New York, New York 10018, 866-364-2733, www.fogcreek.com/FogBUGZ.

Escalade 8506-MI SATA Controllers

The Escalade 8506 Series Multi-lane Internal (MI) Connector RAID controllers are an integrated connector system that combines four SATA ports into a single connection on the controller side and a single connector on the backplane. The Multi-lane Controllers support up to 12 Serial ATA drives on a single PCI card, enabling up to 3TB of storage on a half-length card (dependent on drive capacity). Employing Escalade's StorSwitch switched architecture, the 8506-MI enables Serial ATA's point-to-point architecture performance of up to 1.5GB/sec per port. The MI Controller offers RAID 0, 1, 10, 5 and JBOD support and is 64-bit/66MHz PCI-compliant.

CONTACT 3ware, Inc., 455 West Maude Avenue, Sunnyvale, California 94085, 877-883-9273, www.3ware.com.

Sun Java Desktop on TALIN Notebook

Tadpole Computers announced that it now is offering the Sun Java Desktop System on Tadpole's newest family of notebooks, the TALIN series. Tadpole notebooks offer users easy migration, integrated security mechanisms and the ability to maintain existing UNIX, Java and Linux applications while interoperating with office documents and back-end services. The Java Desktop System includes GNOME, StarOffice, Mozilla, Evolution, Java 2 Platform and a Linux OS. TALIN 15, now available, has a P4 processor with speeds of up to 3.0GHz, 128MB to 1GB of SDRAM, a 15" SXGA+ screen, 3-D graphics support and integrated Wi-Fi. Other TALINs to be released include the 100X, weighing 3.5 pounds, and a 17" screen model.

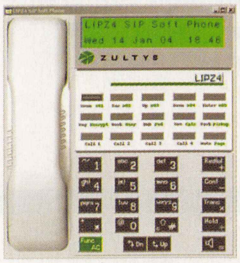
CONTACT Tadpole Computers, Inc., 20245 Stevens Creek Boulevard, Cupertino, California 95014, 800-734-5483, www.tadpolecomputer.com.

Roku HDTV SDK

Roku, maker of the HD1000 high-definition digital media player, recently released a software developer's kit (SDK) for the creation of media applications for high-definition television. The Roku SDK allows developers to create applications for the Roku HD1000 in the C and C++ languages. Developers can use the SDK to access television-centric user-interface elements, media streaming, network and memory card access and device control. In addition, custom installers can control the Roku HD1000 without the SDK by using simple ASCII control commands. The HD1000 is completely controllable from the serial port, Ethernet port or scripts running on the device. The SDK is available for download at www.rokulabs.com/developers.

CONTACT Roku, 399 Sherman Avenue, Suite 12, Palo Alto, California 94306, 866-400-7658, www.rokulabs.com.

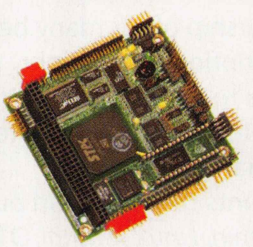
LIPZ4 Soft Phone



Zultys Technologies' LIPZ4 is a freely available software phone that allows users to make and receive phone calls from their computers, without the need for telephone handsets. Based on open standards, the LIPZ4 is compatible with any IP telephony system using Session Initiation Protocol (SIP). It supports four call appearances, instant messaging, hold, transfer, forward, redial and many other features. The LIPZ4 can store the last 32 incoming and outgoing phone numbers. Additional features, including conferencing, compression (G.729) and backup server specification, are available by purchasing a license from Zultys. The advanced encryption standard (AES) is used for network security. LIPZ4 can be downloaded from www.lipz4.com.

CONTACT Zultys Technologies, 771 Vaqueros Avenue, Sunnyvale, California 94085, 408-328-0450, www.zultys.com.

CoreModule 420 PC/104 SBC



Ampro's CoreModule 420 PC/104 SBC is a migration product designed to keep 486-based embedded systems in production through 2010. The I/O addresses and IRQ mappings allow it to be configured to match many legacy system setups. In addition, the memory hole is configurable for legacy PC/104 peripheral cards. The CoreModule 420 supports four serial ports (two for RS-485), along with eight general-purpose I/O (GPIO) pins, an EPP/ECP parallel port, one USB 1.1 port and a high-resolution 2-D video controller with CRT. Interfaces are provided for PS/2 keyboard and mouse, floppy and Ultra/DMA 33 IDE disks, 10/100BaseT Ethernet and TFT flat panels. CoreModule 420 provides 64MB of soldered SDRAM, along with a byte-wide socket for a DiskOnChip 2000 bootable Flash drive and a Type II CompactFlash socket.

CONTACT Ampro Computers, Inc., 5215 Hellyer Avenue, Suite 110, San Jose, California 95138, 800-966-5200, www.ampro.com.

Please send information about releases of Linux-related products to Heather Mead at newproducts@ssc.com or New Products c/o *Linux Journal*, PO Box 55549, Seattle, WA 98155-0549, or fax to 206-782-7191. Submissions are edited for length and content.

LETTERS CONTINUED FROM PAGE 8

Doc is on to the scoop of his journalistic career with his series on how open source (and LAMP in particular) is taking over the enterprise and putting customers back in control. The game is no longer about getting spoon-fed solutions that take too long to implement, cost too much money and fail to address the specific needs of any given business. It's time to say good-bye to cookie-cutter enterprise software and Doc knows it.

What is more, another reason why he is on to something big is because the open-source way, and the LAMP platform in particular, make developers more productive. Vendors know it, developers who have used the tools know it, and even those that haven't are beginning to suspect it.

In my day-to-day, I deal with friends and colleagues (as well as my own staff) whose employers everyone would recognize as being pillars of the Internet economy. These folks depend daily on the huge productivity improvements that come from open development tools and the passionate communities willing to stand behind them. However, most of them either won't, or can't for various corporate reasons, climb the nearest tree and scream to the world about it.

Like dark matter that can't be seen but whose presence is felt everywhere, all of these folks "in the know" are rebuilding the very infrastructure on which IT depends, one brick at a time. And, in perhaps what is just as important, they are also rebuilding all of the starting assumptions behind what "enterprise software" is all about. So thanks Doc and *LJ*. Keep on it—you're on to something big.

--
Antonio Rodriguez, VP Engineering

Errata

In the article "DIY-IT: How Linux and Open Source Are Bringing Do-It-Yourself to Information Technology" [LJ, February 2004], Craig McLane, VP of Technology at Ticketmaster, was mistakenly identified as the source of quotations that were in fact made by Sean Moriarty, EVP of Products and Technology at Ticketmaster. The article also said the event where the talk took place was LinuxWorld Expo. In fact, both men spoke at the O'Reilly Open Source Convention.—Doc Searls

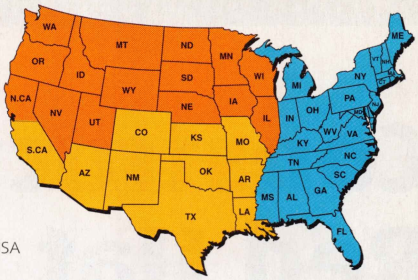
We welcome your letters. Please submit "Letters to the Editor" to ljeditor@ssc.com or SSC/Editorial, PO Box 55549, Seattle, WA 98155-0549 USA.

TECH TIP

If you liked Lisa Corsetti's article "Controlling Hardware with ioctls", in the January 2004 issue, try a more featureful version of the same idea. Run `ifplugd`, which watches all your network interfaces and automatically brings up networking when the cable gets plugged in or the wireless antenna gets a signal.

Get it from your distribution or at **www.stud.uni-hamburg.de/users/lennart/projects/ifplugd**.

LINUX JOURNAL



PO Box 55549
Seattle, WA 98155-0549 USA
www.linuxjournal.com

ADVERTISING SERVICES

VP OF SALES AND MARKETING
Carlie Fairchild, carlie@ssc.com
+1 206-297-8655,
+1 206-782-7191 FAX

FOR GENERAL AD INQUIRIES
e-mail ads@ssc.com
or see www.linuxjournal.com/advertising

To receive further information from any of our advertisers, please visit our Reader Service page to access this service: **www.linuxjournal.com/advertising/rs**

REGIONAL ADVERTISING SALES

NORTHERN USA
Joseph Krack, joseph@ssc.com
866-423-7722 (toll-free),
866-423-7722 FAX

SOUTHERN USA
Annie Tiemann, annie@ssc.com
866-965-6646 (toll-free),
866-422-2027 FAX

EASTERN USA AND CANADA
Martin Seto, mseto@ssc.com
+1 905-947-8846,
+1 905-947-8849 FAX

Please direct international advertising inquiries to VP of Sales and Marketing, Carlie Fairchild.

Advertiser	Page #	Advertiser	Page #
ABERDEEN, LLC www.aberdeeninc.com	91	LINUXCERTIFIED, INC. www.linuxcertified.com	85
AMERICAN MICROSYSTEMS, LTD. www.amltd.com	94	LINUX CLUSTERS INSTITUTE www.linuxclustersinstitute.org	37
APPRO HPC SOLUTIONS appro.com	63, 83	LINUX JOURNAL www.linuxjournal.com	39, 87
ARKEIA CORPORATION www.arkeia.com/demo	19	LINUX NETWORX www.linuxnetworx.com	49
ASA COMPUTERS www.asacomputers.com	35	MICROWAY, INC. www.microway.com	59, C4
ASPEN SYSTEMS, INC. www.aspsys.com	7	NETWORLD + INTEROP LAS VEGAS 2004 lasvegas2004	73
CENTRAL COMMAND, INC. www.centralcommand.com	90	NEX COMPUTERS, INC www.nexcom.com	31
COMPUTER ASSOCIATES (CA) www.ca.com/linux	1	PARACEL www.paracel.com	5
COMTROL CORPORATION www.comtrol.com	25	PARASOFT CORPORATION www.parasoft.com/linux7	69
COYOTE POINT SYSTEMS www.coyotepoint.com	75	PENGUIN COMPUTING www.penguincomputing.com	29
CYCLADES CORPORATION www.cyclades.com	2, 27	POLYWELL COMPUTERS, INC. www.polywell.com	C3
DESKTOPLINUX.COM www.desktoplinux.com	34	THE PORTLAND GROUP www.pgroup.com	43
DICE www.dice.com	89	PORTLOCK SOFTWARE www.portlocksoftware.com	93
EMAC, INC. www.emacinc.com	78	PROGRAMMER'S PARADISE INC. www.programmersparadise.com	65
EMPERORLINUX www.emperorlinux.com	15	PROMICRO SYSTEMS www.promicrosystems.com	55
ETNUS www.etnus.com	61	QSOL.COM www.qsol.com	11
FAIRCOM CORPORATION www.faircom.com	51	RACKSAVER www.racksaver.com	C2, 45
FOURTH GENERATION SOFTWARE SOLUTIONS www.fourthgeneration.com	42	RACKSPACE MANAGED HOSTING www.rackspace.com	13
GOOGLE www.google.com/lj	67	REAL WORLD LINUX www.realworldlinux.com	71
HENTZENWERKE www.hentzenwerke.com	23	SBE, INC. www.sbei.com	9, 21
HURRICANE ELECTRIC www.he.net	77	SGI www.sgi.com	33
IMAGESTREAM INTERNET SOLUTIONS www.imagestream.com	41	SWELL SOFTWARE, INC. www.swellsoftware.com	44
INTEGRATED COMPUTER SOLUTIONS, INC. www.ics.com	22, 84	TATUNG SCIENCE & TECHNOLOGY www.tsti.com	17
IRON SYSTEMS, INC. www.ironsystems.com	86	TECHNOLOGIC SYSTEMS www.embeddedx86.com	47
LANTECH PC SYSTEMS www.lantechpc.com	57	USENIX www.usenix.org/events	79
LAYER 42 NETWORKS www.layer42.net	53		

Constructing Red Hat Enterprise Linux v. 3

Putting together a Linux distribution gets a lot more complicated when stacks of requirements start arriving from hardware vendors and other partners.

BY TIM BURKE

Many people have little understanding of the behind-the-scenes efforts required to construct an enterprise Linux distribution. The fact that this process largely can be taken for granted is actually a compliment. This article offers a glimpse into the methodology we used to deliver Red Hat Enterprise Linux v. 3 (Enterprise Linux v. 3). As you will see, we faced numerous challenges along the way. Then again, if it would have been easy, it wouldn't have been so much fun.

Throughout this article, the focus is on how the release was put together. This article primarily discusses the development of the kernel used in Enterprise Linux v. 3. The kernel is only a fraction of an overall distribution, the portion that controls the underlying hardware and system resources. The challenges faced by the other teams, with projects such as compiler tools, the installer, hundreds of application packages, documentation and testing, are equally daunting. Each of these items was developed by gifted individuals.

Allow me to start by noting that we here at Red Hat have established strong relationships with our key partners in the industry. Although partners' anonymity is guarded here, I'm sure they can identify where they fit in to this story and recognize that it's all in good humor.

What Is an Enterprise Distribution?

High customer expectations have been set by the proprietary UNIX operating systems, and customers planning to migrate from UNIX to Linux do not want to adopt technology that cannot deliver the same level of robustness, quality, support and compatibility. Business users demand stability and reliability. In some cases this means bleeding-edge technology is not appropriate for inclusion in a product. Users also want the ability to run across a wide range of architectures and hardware components, thereby realizing the Linux goals of avoiding proprietary vendor lock-in. Support needs to be in the form of ongoing maintenance for several years, including security and bug fixes, as well as incremental hardware support and valuable, but not destabilizing, feature enhancements.

Requirement Gathering

The ball gets rolling on a new release by identifying the targeted feature set. Next to strong opinions, the second most plentiful commodity at Red Hat is feature requests. They come from all directions, such as insatiable independent hardware/software vendors, demanding customers, both large and small, and Red Hat's worldwide sales and service support organizations. Additionally, as many ideas are generated within Red Hat engineering, ranging from performance and usability enhancements to marketing proposals on how to organize the product set.

Obviously, we don't have infinite developer resources, so the major challenge is to choose the best of these ideas. Another key feature acceptance criteria is conformance with upstream Linux direction, which is necessary for compatibility and to remain true to the spirit of the kernel.org tree governed by Linus Torvalds and containing contributions from around the world.

Here's a few examples of the challenging scenarios we face in requirements gathering:

- We ask each of our partners to submit a reasonably sized top-ten list. One partner's requirements came in the form of a binder that was two inches thick. This became affectionately known as the bible. My first exposure to the bible came when it was heaved onto my desk with a resounding thud. When I saw that, I swear that my heart went thud, too.

- The more mathematically inclined partners do internalize the concept of a top-ten list. However, most ended up using a tactic that those of you familiar with TCP/IP networking should recognize—the sliding window protocol. The way it works is as soon as any of your features have been accepted, those pop off the top of the stack, freeing up space for features 11, 12 and 13 to all of a sudden become cataclysmic issues.

The following is a representative example feature list from a hardware vendor:

1. Support more than 32GB of memory on x86.

2. Support more than eight CPUs.

3. Support for our new XYZ100 series computers. (This is a thinly veiled multiple feature request; behind it is a series of required device drivers, PCI IDs and installer hooks.)

4. Updated I/O adapter driver. (This ultimately turns into differences of opinions regarding to which newer version this refers.)

5. Integrated support for the vendor's proprietary baseboard management software. (A perennial list item, which consistently gets rejected to the amazement of the requester.)

6. Compiler optimizations to match the latest chipsets.

7. USB support for our CD-ROM drive. (Needed because the drive is brain-dead and doesn't conform to the spec—of

course that subtlety is absent in the initial feature request.)

8. Support for more than 128 disks.

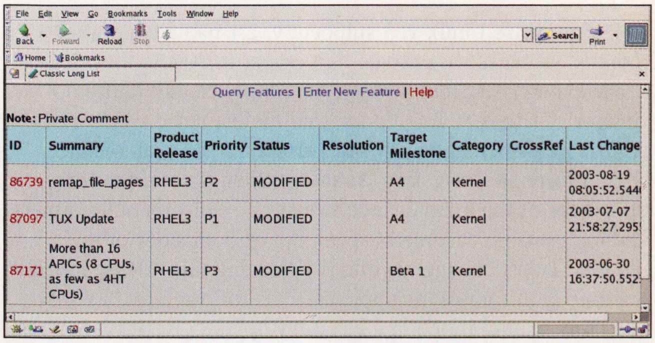
Then, there's typically the following implied requirements:

- All of these feature requests apply to multiple architectures, including x86, AMD64 and Itanium 2.
- Oh, by the way, we also want this feature backported to the prior Enterprise Linux v. 2.1 release.

Countless external requests ask for either proprietary additions or hooks. In the open-source tradition, this is something we consistently have to refuse.

For Enterprise Linux v. 3, the initial set of requested features was about 500 items. Each of these was entered into our feature-tracking tool called Featurezilla, which actually is a derivative of the Bugzilla bug-tracking tool. Figure 1 shows several items in Featurezilla. To show that the jabs go both ways, one partner has a business manager whose name is Paul, name changed to protect the innocent. He manages a list as a text file, but back at his company, the list is referred to as Paulzilla.

We had countless tortuous internal meetings to prioritize and slog through the full set of 500 items. From that prioritized list, the engineering managers went off to try to get the list down to a manageable set that is humanly achievable by their team members. Ultimately, the list is negotiated down to something that vastly exceeds anything reasonably achievable by the development team. But, that's the Red Hat way, and it's only the beginning of the story.


File Edit View Go Bookmarks Tools Window Help
Back
Forward
Reload
Stop
Search
Print
Home
Bookmarks
Classic Long List
Query Features | Enter New Feature | Help

Note: Private Comment

ID	Summary	Product Release	Priority	Status	Resolution	Target Milestone	Category	CrossRef	Last Change
86739	remap_file_pages	RHEL3	P2	MODIFIED		A4	Kernel		2003-08-19 08:05:52.544
87097	TUX Update	RHEL3	P1	MODIFIED		A4	Kernel		2003-07-07 21:58:27.295
87171	More than 16 APICs (8 CPUs, as few as 4HT CPUs)	RHEL3	P3	MODIFIED		Beta 1	Kernel		2003-06-30 16:37:50.552

Figure 1. The feature-tracking tool Featurezilla lists feature requests and their status.

Red Hat Enterprise Linux v. 2.1 Maintenance Pulls

To keep things interesting, in addition to having to develop to an aggressive schedule for Red Hat Enterprise Linux v. 3 development, a large amount of effort also is required to support the Enterprise Linux v. 2.1 maintenance stream. By the

Appro High-Performance Workstations, Servers and Cluster Solutions

Appro servers and workstations based on AMD Opteron™ processors boost performance on existing Windows® and Linux applications.

Appro WH300 Workstation

Dual AMD Opteron™ Processors
- Designed for demanding compute-intensive applications
- Delivers 8x AGP graphics capabilities for content creation, animation, 3D modeling and CAD to meet all your design and analysis needs
- Provides robust memory and storage with up to 16GB of DDR333/400 RAM
- Integrated FDD and CD/DVD drive
- Supports ATI®, NVIDIA and other graphics solutions

Appro 1U 1122H Servers

Dual AMD Opteron™ Processors
- Design to run cool & efficiently in demanding enterprise environments
- Ideal for super-computing clusters
- Up to 16GB of DDR 333/400 RAM
- Up to 4 HDDs - IDE or SCSI
- Dual Gigabit ports
- 24x slim CD-ROM
- Integrated Cluster Solutions are available

Appro 4U 4145H Server

New!

Quad AMD Opteron™ Processors
- Provides performance and capability for high-end graphic simulations
- Up to 32GB in 16 DIMM sockets
- Up to 4 SATA or SCSI HDDs
- 5.25" bay for CD-ROM or tape devices
- Dual Gigabit ports
- 3.5" floppy drive

Appro 2U Servers

Dual AMD Opteron™ Processors
- Delivers high performance with expandable storage capability
- Up to 16GB of DDR 333/400 RAM
- Up to 8 HDDs – IDE, SATA or SCSI
- Dual Gigabit ports
- 24x slim CD-ROM drive

AMD Opteron™ Processors - Shorten run-time cycles and increase bandwidth for processing computing requests.
- 32 bit applications while you migrate to 64 bit computing for long-term investment protection.

For more information, please visit www.appro.com
or call Appro Sales at 800.927.5464, 408.941.8100.

APPRO HPC Cluster Solutions

AMD 64 Opteron

time Enterprise Linux v. 3 shipped, v. 2.1 had been out for 15 months. Because we provide a five-year life cycle for releases in the Enterprise Linux family, we simply cannot dump Enterprise Linux v. 2.1 support and maintenance work. The requirements demanded by our partners in the maintenance stream form an interesting paradox. When it comes to the objectives of the maintenance stream, it seems that all partners and high-profile customers speak out of both sides of their mouth. "Don't change anything! I'm using this release in production. Don't upset the apple cart", is followed by "I really need this feature immediately. Hurry up and give it to me, but don't put anything else in."

Ultimately, it all boils down to a careful case-by-case risk/benefit assessment. Internally, we review features and bugfix proposals among the engineering ranks. It is well beyond my literary abilities to convey the strong-willed and passionate debates conducted over incorporating features or bug fixes that straddle the risk/benefit fence.

Red Hat Enterprise Linux v. 3 Kernel Development

Some people have the mistaken impression that Red Hat simply pulls together a random collection of bits and pieces from the upstream Open Source community, slaps the Red Hat name on it and calls it a product. Truth be told, Red Hat engineers contribute a substantial percentage of upstream (for example, 2.4 and 2.6) kernel development. The productivity, breadth of knowledge and ability to be on top of a torrent of internal and external communication demonstrated by these kernel developers is stunning. They truly are world class and humbling to be among. But don't tell them that, lest it goes to their heads.

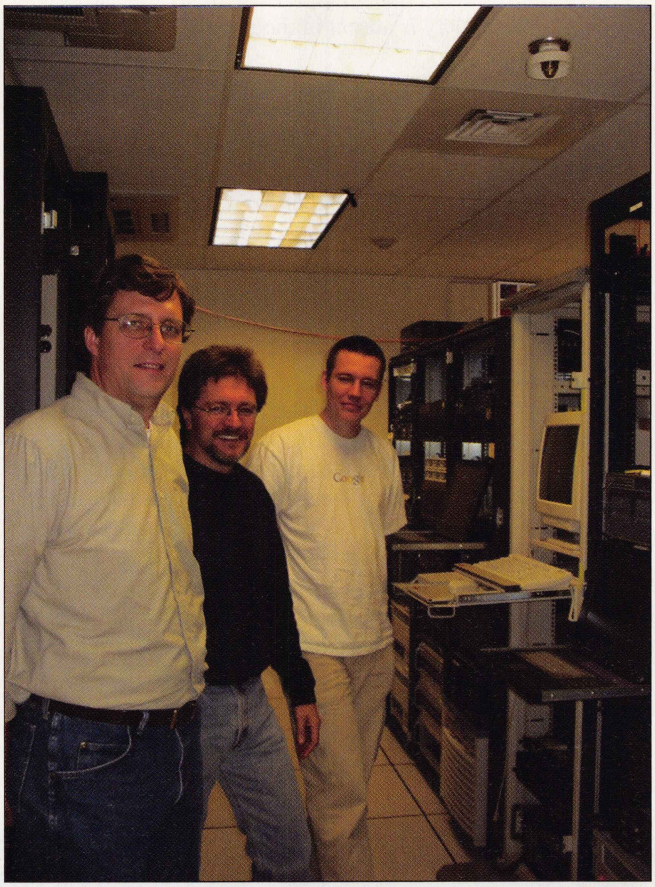


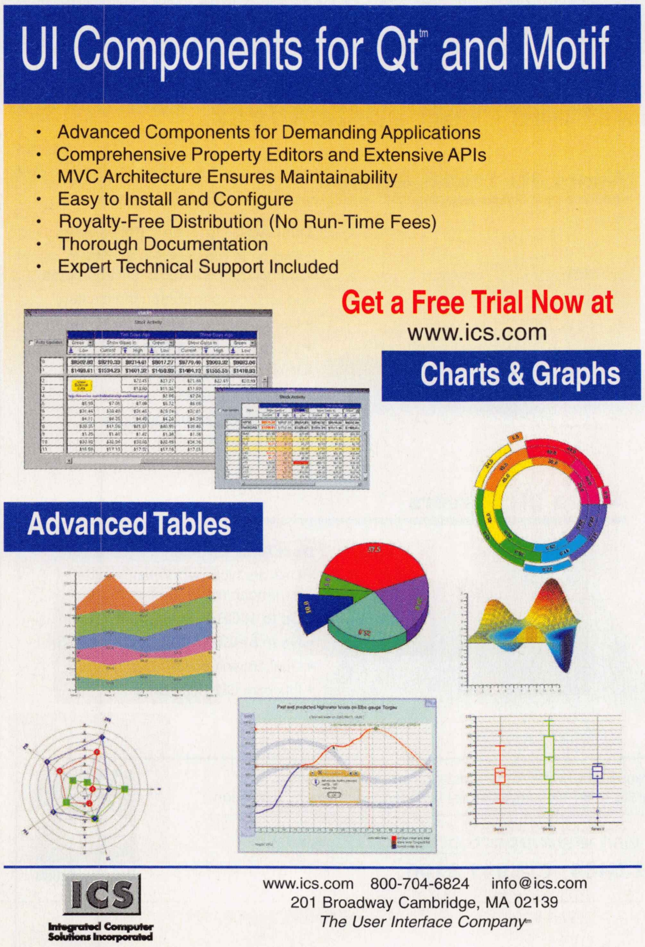
Figure 2. Kernel developers at Red Hat, left to right: Larry Woodman, Dave Anderson and Rik van Riel.

In addition to the sizable upstream enhancements that get pulled back into the Enterprise Linux kernel, a large set of enhancements is developed in-house to meet the product requirements. As an open-source company, all of the kernel enhancements are available to the community at large. The vast majority of these changes do end up being incorporated upstream.

The kernel in Enterprise Linux v. 3 primarily is based on the 2.4.21 kernel, but it has a huge number of features backported from the more recent 2.6 kernel. In recognition that the 2.6 kernel had not yet been stabilized, only key features deemed commercially ready were candidates for the backport into Enterprise Linux v. 3. Here's a little tip on how to really annoy a Red Hat kernel developer: say something like, "So, you guys just ship the stock 2.4.21 kernel; right?"

The most daunting challenges in constructing the Enterprise Linux v. 3 kernel were the requirements that support be provided for seven different architectures and that the kernels for these architectures all must be built from a common source tree. The seven architectures include: x86 (and Athlon), AMD64, Itanium 2, IBM mainframe (both s390 and s390x) and IBM's iSeries and pSeries PPC systems. Although each of these architectures is supported by the upstream kernels in varying degrees, the reality is many of these architectures are developed primarily in isolation. This inevitably leads to integration nightmares.

Another interesting twist to the story is Red Hat's kernel devel-



UI Components for Qt™ and Motif

- Advanced Components for Demanding Applications
- Comprehensive Property Editors and Extensive APIs
- MVC Architecture Ensures Maintainability
- Easy to Install and Configure
- Royalty-Free Distribution (No Run-Time Fees)
- Thorough Documentation
- Expert Technical Support Included

Get a Free Trial Now at
www.ics.com

Charts & Graphs

Advanced Tables

ICS
Integrated Computer Solutions Incorporated

www.ics.com 800-704-6824 info@ics.com
201 Broadway Cambridge, MA 02139
The User Interface Company™

opment team literally is spread all over the world. Out of necessity, this has led to virtually all interaction being done on-line. For example, we rarely have team meetings, as time-zone challenges get in the way. Sometimes this on-line communication goes to extremes; it's quite common to use IRC chat sessions to speak with someone sitting at the next desk. Yes, our mothers were right; we are a bunch of geeks.

Late-Breaking Features

Just as the sun never sets on kernel development, our near-and-dear partners and customers don't sit still either. One of the few constants in this dynamic environment is that throughout the course of the release development, there is a steady influx of must-have crisis feature additions. We endeavor to make it a trade-off and kick out a lower priority feature in order to keep the workload sane. In the end, however, it never seems to work out to such a sweet balance. The ability of the team to persevere through all these demands is remarkable.

As productive as the Red Hat developers are, some late-breaking features always have to be deferred to a later release or update. These situations cause no end of trauma for our partner managers who have to be the bearers of bad news. It's improbable that our partners ever heard the saying "don't shoot the messenger". There was one incident when we were two hours from shipping the release and a delivery arrived on the loading dock. It was a new computer platform we needed in order to be able to develop and test support for it. The partner was incredulous that we were unable to accommodate.

Testing

Several different levels of testing are performed throughout the development process. It all begins with the developers performing unit testing. This consists of manual, hands-on testing as well as development of automated testing programs. The set of automated tests constantly is being augmented as new problems are addressed. These automated test programs then are incorporated into a test grid that is managed by the QA department. In addition to the internally written unit tests, our test grid includes a wide range of regression tests. Examples include POSIX, LSB conformance, LTP, crashme, gcc suite and diabolical tests provided by our partners. Stress tests also are performed for a range of system functions using such tests as cerberus, lmbench, bonnie, spec and other microbenchmarks, to name a few. These automated tests are run nightly in order to detect regressions quickly. This is critical in order to isolate the offending code. In contrast, if we waited to perform monthly base-level testing, it would be much more difficult to identify the culprit.

On top of the nightly tests, more time-consuming test scenarios are run less frequently. Examples include installation testing using a wide range of configuration options across the many different languages we support. Hands-on testing rounds out the internal QA coverage. Doing justice to the staggering range of testing done by the hardworking QA team here substantially exceeds the scope of a single article.

After the kernels have passed internal units tests and QA scrutiny, we make them available to our development partners. This vastly broadens the coverage to include external QA and development teams in other companies. Our partners focus testing on their hardware platforms as well as the typical server capabilities their enterprise customers demand.

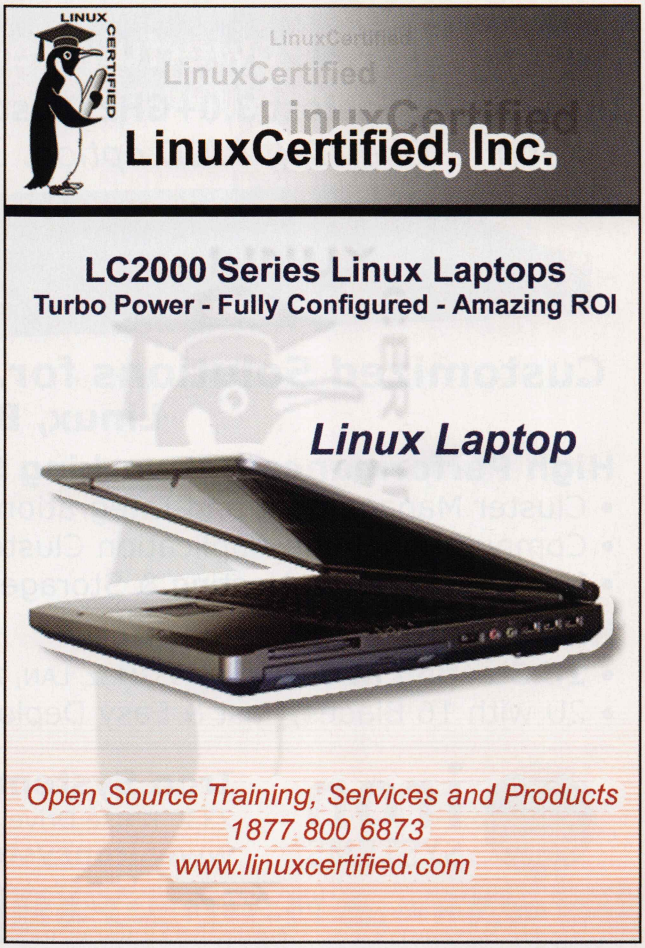
The last layer of testing includes external beta testers. These beta testers include high-profile customers, as well as the many people in the Linux community who respond to our open invitation to help with testing.

The combination of all these different testing activities yields an extremely stable and well-tested product. It also points out the huge value of the open-source model. It is the combination of testing resources from many different companies and dedicated individuals that scales well beyond the resources a single company could bring to bear to tackle an infinite testing matrix.

Figure 3, composed by Nick Carr, summarizes the development model from requirements gathering, development and testing, culminating in production.

Conclusion

In the end, the Red Hat Enterprise Linux v. 3 distribution did ship on schedule in mid-October 2003. Having worked for many years for a major IHV on a large-scale operating system engineering team, I have been exposed to both the proprietary operating system development model and the open-source development model. The successful outcome of the Enterprise Linux v. 3 release clearly demonstrates the power of cooperative open-source development. The quantity of features, rapid development time and caliber of participants, both internal and external, is remarkable and



LinuxCertified, Inc.

LC2000 Series Linux Laptops
Turbo Power - Fully Configured - Amazing ROI

Open Source Training, Services and Products
1877 800 6873
www.linuxcertified.com

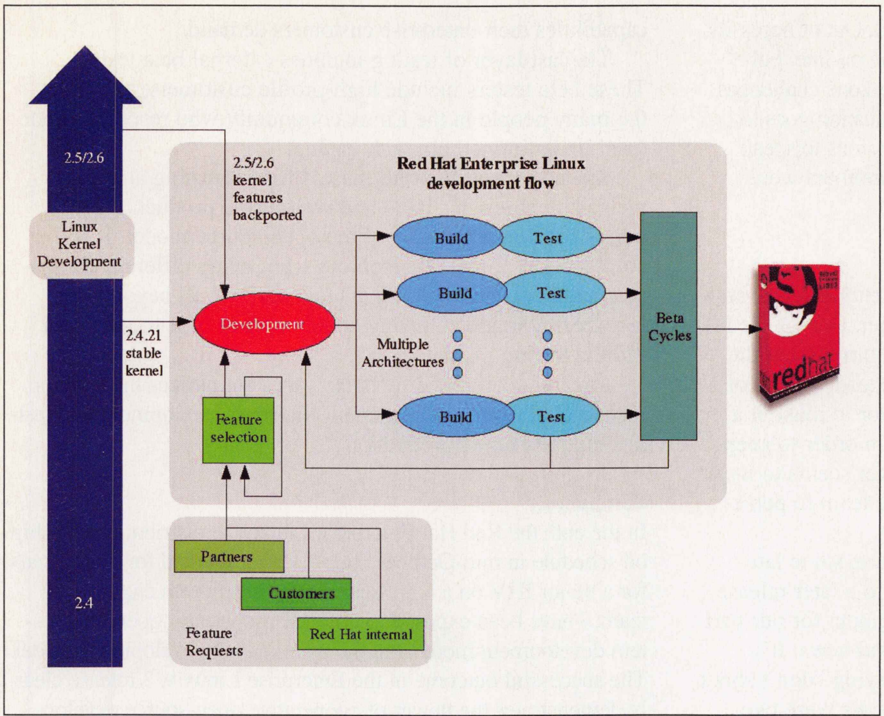
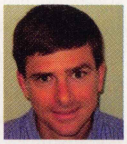

2.5/2.6
Linux Kernel Development
2.4.21 stable kernel
2.4
2.5/2.6 kernel features backported
Red Hat Enterprise Linux development flow
Development
Build
Test
Multiple Architectures
Beta Cycles
Feature selection
Partners
Customers
Red Hat internal
Feature Requests
redhat

Figure 3. The Development Process, from Requirements Gathering to Release

stands as a strong testament to the Linux community.

When the release shipped, we all were proud to have contributed to such a tremendous accomplishment. But the time window to rejoice was short indeed. As soon as Enterprise Linux v. 3 finished, it seemed we already were late with the development of Enterprise Linux v. 4. This reminds me of the quote, "Congratulations for winning this battle, that earns you the privilege to come back and fight another day." Gotta go, more battles to fight.

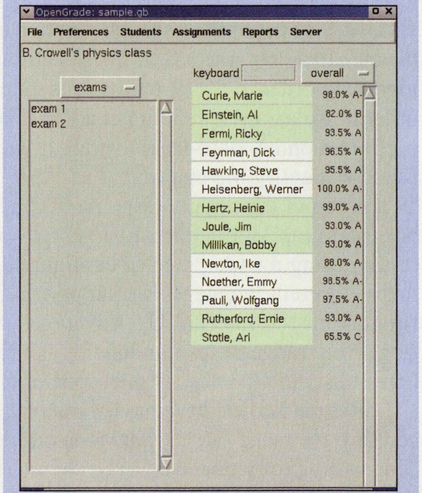


Tim Burke is the director of Server Development at Red Hat. This team is responsible for the kernel and the set of core applications included in Red Hat Enterprise Linux. Prior to becoming a manager, Tim earned an honest living developing Linux highly available cluster solutions and UNIX kernel technology.

Ultra Dense, Powerful, Reliable...
Introducing Fastest 3.0+GHz Clustering Node!
2-Xeon, 2-Gbe LAN, RAID option, 15" deep
Customized Solutions for...
Linux, BSD, W2K
High Performance Networking Solutions
• Cluster Management and Integration
• Computational and Application Clustering
• Network Security, Hosting & Storage Engines
Rackmount Server Products
• 1U Starting at $499: C3-933MHz, LAN, 256MB, 20GB IDE
• 2U with 16 Blades, Fast & Easy Deployment
iron SYSTEMS™
Iron Systems, Inc.
2330 Kruse Drive, San Jose, CA
www.ironsystems.com
CALL: 1-800-921-IRON

OpenGrade:
lightandmatter.com/ogr/ogr.html

OpenGrade can be used to maintain grades for a class and can calculate totals based on either a point system or weighted averages. Exams, homework and so on are input and scored independently. The system can be told which grading standards to use and more. Reports are quick and easy to generate. Documentation is excellent. Requires: Perl, Perl modules Term::ReadKey and TermUI.


OpenGrade: sample.gb
File
Preferences
Students
Assignments
Reports
Server
keyboard
overall
exams

B. Crowell's physics class

exam 1
exam 2

Curie, Marie	98.0% A-
Einstein, Al	82.0% B
Fermi, Ricky	93.5% A
Feynman, Dick	96.5% A
Hawking, Steve	95.5% A
Heisenberg, Werner	100.0% A-
Hertz, Heinie	99.0% A-
Joule, Jim	93.0% A
Millikan, Bobby	93.0% A
Newton, Ike	88.0% A-
Noether, Emmy	98.5% A-
Pauli, Wolfgang	97.5% A-
Rutherford, Ernie	93.0% A
Stotle, Ari	65.5% C-

—DAVID A. BANDEL

Let the games begin!

You may be running out of pages in your *Linux Journal* calendar, but the contests won't end there! We are pleased to announce that during 2004 we are going to continue to introduce a new contest each month.

Each contest will be listed on the official contest web site: **http://www.linuxjournal.com/contest**, and will also be announced in the *Linux Journal* weekly e-newsletter. (Visit http://www.linuxjournal.com/ to subscribe and make sure you get a head start on each contest!)

You can look forward to more trivia questions, challenges, and exciting prizes, much like these previous contests...

Calling All Admins (September 2003)
We're looking for the remotest administration: send us your latitude and longitude, the latitude and longitude of the furthest-away Linux system you have admin responsibility for, and the distance between them. The person with the furthest distance will win a lifetime subscription to *LJ* and also fun Linux schwag.

Beer Trivia (July 2003)
How many *Linux Journal* covers pictured beer on them? One winner, who will be enrolled in a Beer-of-the-Month club for one year, will be randomly selected from all correct submissions. Beer prize available to residents of the continental USA only and must be 21 years of age or older (all are welcome to enter the contest, if a winner who does not fit these requirements is selected, another prize will be presented).

LINUX
JOURNAL

For more information visit http://www.linuxjournal.com/contest

Samba Logging for Audit Trails

Audit trails are a network security requirement for both Northrop Grumman and its customers. A small modification to Samba enabled the company's sysadmins to create the needed audit trails.

BY EDWARD KABLAOUI

We at Northrop Grumman recently decided to replace the Microsoft Windows 2000 server on one of our networks with a Linux server running Samba. The primary motivations for replacing the Windows server with Linux were:

- Common user names and passwords for both Windows and Linux users.
- Export commonly shared directories and files using NFS and Samba.
- Allow software developers the freedom to choose their development environment.
- No additional licensing fees required for Windows 2000 server and clients.
- Centralized and cleaner audit trails.
- A more secure computing environment.
- Software upgrades can be scheduled when necessary as opposed to being dictated by outside software vendors.

In addition to configuring Samba as a primary domain controller, modifications were made to the Samba source code to meet the security requirements of our network. This article briefly describes how to install and configure Samba and then explains in detail how to modify Samba source code to produce log entries required for audit trails.

Audit Trail Requirements

Our networks must be configured to meet corporate and customer requirements regarding security. Among the various security requirements that must be met, administrators need to have the ability to audit activity on the network. The required information gathered and logged for these audits is referred to as an audit trail and include the following information: successful logins, logouts, failed logins and password changes.

Most operating systems generate log files with this information, but it may be scattered in a number of different files and contain more information than is necessary. We also wanted to centralize this information so network administrators do not have to examine the logs on every computer on the network. Configured properly, a Windows 2000 server logs the above information for all machines connected to the network, in addition to reporting a myriad of extraneous information. This additional information makes security audits longer, more error-prone and also occupies a lot of disk space when archived.

In order to use Linux and Samba as the primary domain server for Windows 2000 clients, Samba has to duplicate the logging capabilities of the Windows 2000 server. Once Samba was configured for the network, it seemed the only way to meet the logging requirements was to modify the Samba source code. An additional benefit of modifying the source code was the ability to have only the necessary information recorded in the log files. The replacement of the Windows 2000 server with Samba would not have been possible had Samba been a closed-source, proprietary product.

Downloading and Configuring Samba

The Linux server initially arrived with Red Hat 8.0 and Samba already installed. The first step was to download the tarred and compressed version of Samba 2.2.8a from the Web site, www.samba.org. Once downloaded, Samba was installed by running the following commands as root:

```
tar cvfz samba-2.2.8a.tar.gz
cd samba-2.2.8a/source
./configure
make
make install
```

The Samba executables, smbd and nmbd, were installed under the /usr/local/samba/bin directory. The Red Hat installation had placed these executables under the /sbin directory. Using the newly created Samba executables requires changing the Samba startup script /etc/init.d/samba. The current Samba dæmons should be stopped first by running the command `/etc/init.d/samba stop`.

The /etc/init.d/samba file then is edited such that the commands for starting smbd and nmbd are changed from `/sbin/smbd -D` to `/usr/local/samba/bin/smbd -D` and `/sbin/nmbd -D` to `/usr/local/samba/bin/nmbd -D`. The new dæmons are then started with the command `/etc/init.d/samba start`.

Once the new dæmons are installed successfully, Samba needs to be configured by setting parameters in the smb.conf file. For the 2.2.8a distribution, the default location of this file is /etc/samba. The smb.conf file consists of sections denoted by square brackets, and each section names a share or service. The following example shows some of the parameter values set under the global section to create a Primary Domain Controller for Windows clients:

```
[global]
netbios name = SambaServer
workgroup     =  NETDOMAIN
domain master = yes
local master = yes
preferred master =yes
os level =  65
```

For authenticating users on the network, the following parameters also need to be set under the global section:

```
encrypt passwords = yes
security = user
domain logons = yes
```

Finally, for Windows 2000 clients, the domain admin group and add user script global parameters need to be set as well:

```
domain admin group = root
add user script = /usr/sbin/useradd -d /dev/null \
                      -g 100 -s /bin/false -M %u
```

Public and private shares for Window clients are created by adding new sections.

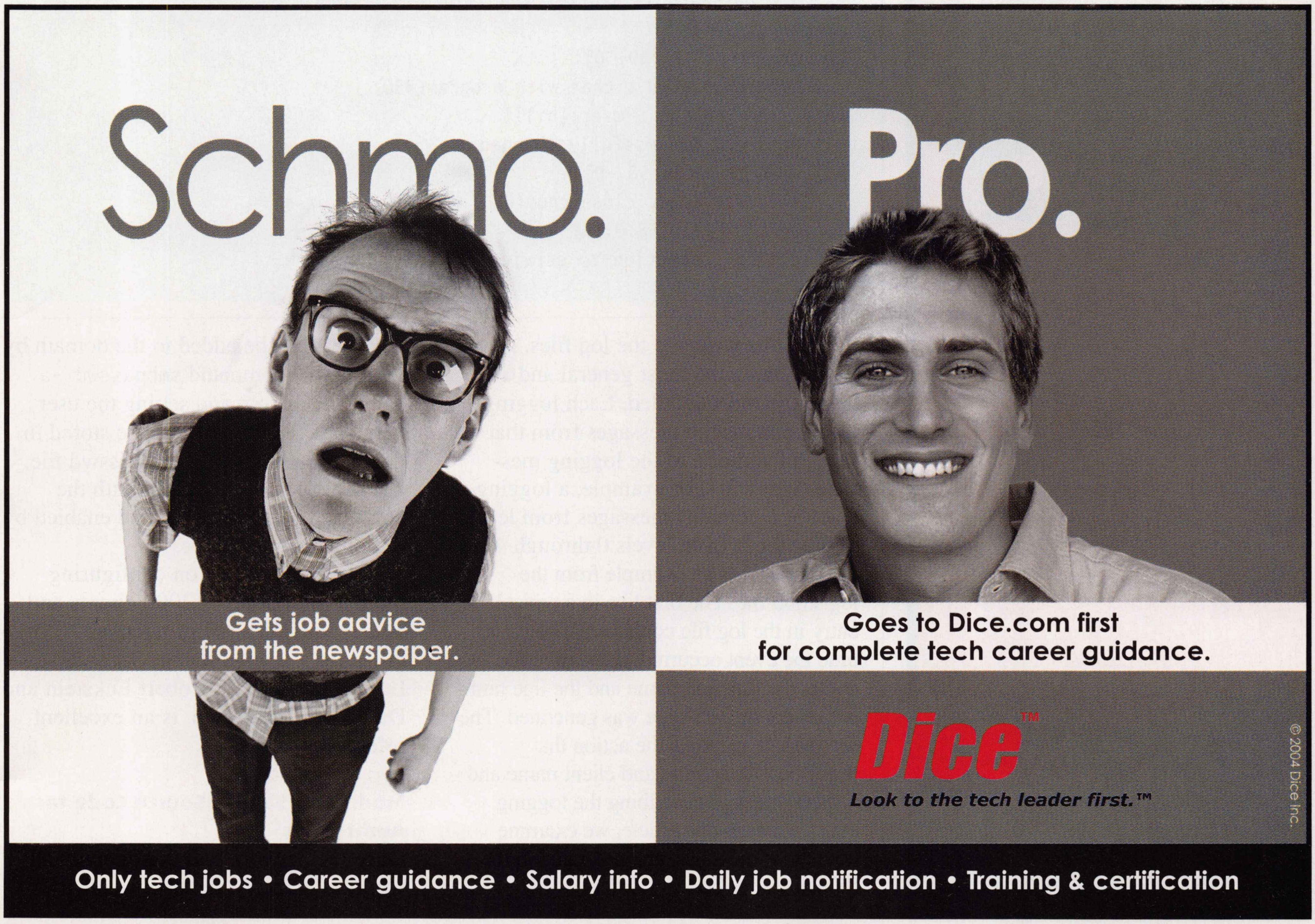
Public:

```
[share]
path = /home/share
read only = no
browseable = yes
guest ok = no
create mode = 0770
comment = Shared Folder
hide dot files = yes
```

Private:

```
[homes]
path = /home/%u
read only = no
browseable = no
guest ok = no
map archive = yes
create mode = 0750
comment = Home Directories
hide dot files = yes
```

To verify that the parameters are correct in the smb.conf file or to debug configuration problems, use the testparm command. For debugging problems with Samba in general, the log files log.smbd and log.nmbd under the /var/log/samba directory are invaluable. The parameter log level in the global section of the smb.conf file determines the amount of detailed informa-


Schmo.
Gets job advice from the newspaper.
Pro.
Goes to Dice.com first for complete tech career guidance.
Dice™
Look to the tech leader first.™
© 2004 Dice Inc.
Only tech jobs • Career guidance • Salary info • Daily job notification • Training & certification

Best of Breed

Virus Protection for Mail Servers

VEXIRA AntiVirus

Linux FreeBSD OpenBSD

VEXIRA Antivirus

Virus Protection for the Real World.

Mail Server

CENTRAL COMMAND

FREE TRIAL!

Vexira™ Antivirus for Mail servers is an industry leading, highly scalable virus scanning solution for email servers. Contact Central Command® today to find out why Vexira Antivirus is known as the "best of breed" in virus protection.

Supported OS:
- Linux
- FreeBSD
- OpenBSD

Email Servers:
- Sendmail
- Postfix
- Qmail
- Exim

Central Command, Inc.
phone: 330.723.2062 • fax: 330.722.6517
www.centralcommand.com

© 2003 Central Command, Inc. All rights reserved. Vexira, Vexira logo, and Central Command are trademarks or registered trademarks of Central Command, Inc. All other trademarks are property of there respective owner(s).

Listing 1. Log Entries from /var/log/samba/log.smbd

```
[2003/11/25 17:13:12, 0] smbd/server.c:main(791)
    smbd version 2.2.8a started.
  Copyright Andrew Tridgell and the \
    Samba Team 1992-2002
[2003/11/25 17:17:32, 0] \
    rpc_server/srv_netlog_nt.c:_net_sam_logon(643)
  Logon . Domain:[NETDOMAIN].\
    HostName:[192.168.0.15]. \
    User:[john]. FAILED No Such User
[2003/11/25 17:17:55, 0] \
    rpc_server/srv_netlog_nt.c:_net_sam_logon(665)
  Logon . Domain:[NETDOMAIN]. \
    HostName:[192.168.0.15]. \
    User:[bill]. FAILED Incorrect Password
[2003/11/25 17:18:33, 0] \
    rpc_server/srv_netlog_nt.c:_net_sam_logon(691)
  Logon . Domain:[NETDOMAIN]. \
    HostName:[192.168.0.15] \
    User:[bill] Successfully Logged On
[2003/11/25 17:19:34, 0] \
    smbd/chgpasswd.c:check_oem_password(836)
  check_oem_password: incorrect password length \
    (262218674) for user bill.
[2003/11/25 17:19:46, 0] \
    smbd/chgpasswd.c:chgpasswd(474)
  Password Change: user bill, \
    New password is shorter than minimum password \
    length = 8
[2003/11/25 17:21:29, 0] \
    smbd/chgpasswd.c:chat_with_program(450)
  Password Change . User:[bill] \
    Password Successfully Changed
[2003/11/25 17:16:58, 0] \
    smbd/service.c:close_cnum(680)
  clientPC (192.168.0.15) \
    closed connection to service bill
```

tion Samba writes to the log files, with level 0 being the most general and 10 being the most detailed. Each logging level contains the messages from that level, in addition to the logging messages below it. For example, a logging level of 5 contains messages from level 5, plus those from levels 0 through 4.

Listing 1 is an example from the log.smbd file. The first line in a typical entry in the log file contains the date and time the event occurred, the source file name, the function name and the line number where the message was generated. The second line contains the action that occurred, the domain and client name and a short message describing the logging event. Later in this article, we examine how these messages are generated in the Samba source code.

Users can be added to the domain by running the command `smbpasswd -a username` as root and setting the user password. The passwords are stored in the /etc/samba/private/smbpasswd file. Users also can be disabled with the command `smbpasswd -d` and enabled by running `smbpasswd -e`.

For more details on configuring Samba for Windows 2000 clients and understanding SMB protocols, O'Reilly's *Using Samba*, Second Edition, by Jay Ts, Robert Eckstein and David Collier-Brown, is an excellent reference.

Modifying Samba Source Code for Audit Trails

Once the network was configured and working properly, the next step was to

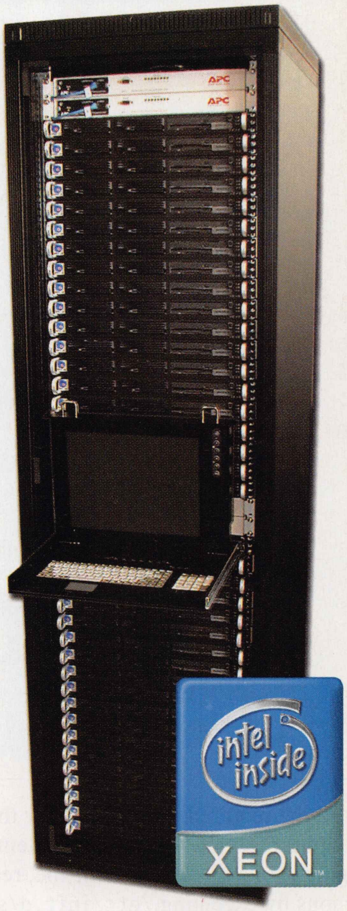
The Straight Talk People℠
SINCE 1991

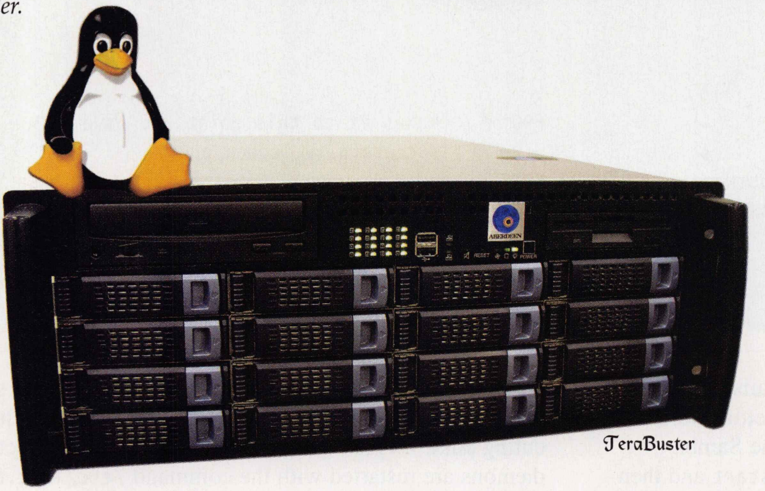
ABERDEEN

4 TERABYTES UNDER $12,000

You're already confidently saving a fortune using the finest operating system available. Why not do the same with your storage needs? After shocking the server storage market with the 2 terabyte Stirling "Backup Monster" 3U solution, Aberdeen introduced the 4 terabyte Stirling "TeraBuster." Packing 16 hot-swappable 250GB SATA drives and triple-redundant 650-Watt power supplies in a 4U package, the "TeraBuster" establishes new standards in price, performance and packaging. Five 8cm fans and compact SATA cabling provide unimpeded airflow for this cool running reliable server. There is even room for two additional internal hard drives to mirror an operating system. All of this at an unheard of price that's under $12,000.

Give us a call and one of our expert system consultants will help you determine which solution best fits your needs or visit our website www.aberdeeninc.com to view more detailed specs of this product. Our industry-leading 5-Year Warranty reduces cost of ownership and provides peace of mind. See why Fortune 500 companies such as Agilent, Boeing, Lockheed Martin, Motorola and Northrop Grumman have chosen Aberdeen, the Custom Server Solution Provider.





TeraBuster

Model	**Stirling X38i "Backup Monster"**	**Stirling X38saR "Backup Monster"**	**Stirling X416R "TeraBuster"**
Form Factor	3U Rackmount	3U Rackmount	4U Rackmount
Intel® Processors	Xeon® 2.4BGHz	Dual Xeon® 3.06GHz	Dual Xeon® 3.06GHz
Chipset	ServerWorks GC-SL	ServerWorks GC-SL	Intel E7501
ECC Registered Memory	1GB (max 8GB) DDR266	2GB (max 8GB) DDR266	2GB (max 16GB) DDR266
RAID Controller	3ware® Escalade™ 7506	3ware Escalade™ 8506	2 x 3ware® Escalade™ 8506
Hot Swap Drive Enclosures	8 x IDE Included	8 x SATA Included	16 x SATA Included
Additional External Drive Bays	2 x 5¼ (free), 1 x 3½, 1 slim CD	2 x 5¼ (free), 1 x 3½, 1 slim CD	1 x 5¼, 1 x 3½
Included Hard Drives	8 x 250GB 5400rpm ATA/133 2MB cache	8 x 250GB 7200rpm SATA/150 8MB cache	16 x 250GB 7200rpm SATA/150 8MB cache
Integrated Network Controllers	2 x Gigabit	2 x Gigabit	2 x Gigabit
AGP VGA Controller	ATI® Rage XL 8MB	ATI Rage XL 8MB	ATI Rage XL 8MB
Available PCI Slots	1 x 64-bit, 3 x 32-bit	2 x 64-bit, 2 x 32-bit	4 x 64-bit
1.44 MB Floppy Disk Drive	Included	Included	Included
CD-ROM	32X Low Profile	32X Low Profile	52X
Tape Drive SCSI	Optional	Optional	Optional
Cooling Subsystem	3 x 9 cm Fans	3 x 9 cm Fans	5 x 8 cm Fans
Power Supply	460 Watt	460 Watt Redundant	650 Watt Triple Redundant
Operating System	Optional	Optional	Optional
Backup Software	Optional	Optional	Optional
Limited Warranty	***5-Year***	***5-Year***	***5-Year***
Price	***$4,465***	***$5,865***	***$11,695***

Intel, the Intel Inside Logo and Intel Xeon are registered trademarks of Intel Corporation. lj001

888-297-7409
www.aberdeeninc.com/linux

Listing 2. DEBUG Calls Added to source/rpc_server/srv_netlog_nt.c

```
632  /* get the account information */
633  pdb_init_sam(&sampass);
634  become_root();
635  ret = pdb_getsampwnam(sampass, nt_username);
636  unbecome_root();
637
638
639  if (!ret)
640  {
641      pdb_free_sam(sampass);
642      // ESK
631  /* checks for failed users */
643      DEBUG(0,("Logon . Domain:[%s]. HostName: \
         [%s]. User:[%s]. FAILED No Such User \n",
                 lp_workgroup(),
                 client_addr(),
                 nt_username));
644      return NT_STATUS_NO_SUCH_USER;
645  }
646
647  acct_ctrl = pdb_get_acct_ctrl(sampass);
630  map_username(nt_username);
.
.
.
663  /* Check for failed password */
664  if (!NT_STATUS_IS_OK(status)) {
665      DEBUG(0,("Logon . Domain:[%s]. HostName: \
    [%s]. User:[%s]. FAILED Incorrect Password \n",
                 lp_workgroup(),
                 client_addr(),
                 nt_username));
666      pdb_free_sam(sampass);
667      return status;
668  }
.
.
.
669 /* Check PAM Password */
670 #ifdef WITH_PAM
671  become_root();
672  status =
    smb_pam_accountcheck(pdb_get_username(sampass));
673  unbecome_root();
674  if (!NT_STATUS_IS_OK(status)) {
675      pdb_free_sam(sampass);
676      DEBUG(0,("Logon . Domain:[%s]. HostName: \
    [%s]. User:[%s]. FAILED Incorrect Password \n",
                 lp_workgroup(),
                 client_addr(),
                 nt_username));
677      return status;
678  }
679 #endif
.
.
.
690    /* Makes it to this point you have
          successfully logged on */
691    DEBUG(0,("Logon . Domain:[%s]. HostName: \
        [%s] User:[%s] Successfully Logged On\n",
                 lp_workgroup(),
                 client_addr(),
                 nt_username));
```

record the required information for the audit trails in the log.smbd file. Initially, this was attempted by setting various logging levels in the smb.conf file, restarting the Samba dæmons by executing `/etc/init.d/samba restart` and then analyzing the output from the log files while performing various tasks on a Windows 2000 client. Unfortunately, regardless of the log level, none of the required information was being logged. At this point, it became obvious that modifications to the source code were necessary to generate the log entries.

The Samba source code is in the samba-2.2.8a/source directory and is the root directory for the filename in the log messages. The first entry in Listing 1 shows the file as `smbd/server.c:main(791)`. By examining line 791 in the samba-2.2.8a/source/smbd/server.c file, you can see the DEBUG macro is used to generate the log message. The syntax for the DEBUG macro is:

```
DEBUG(log_level, "string", arguments);
```

The format for the second and third arguments is similar to the printf function. Examples of custom DEBUG statements are shown in Listings 2 through 4.

Creating custom log messages requires inserting DEBUG macros into the appropriate section of code and filling in the correct parameters and messages. Once the DEBUG statements are inserted, the Samba executables need to be rebuilt by executing `make install` in the samba-2.2.a/source directory; the dæmons are restarted with the command `/etc/init.d/samba restart`. Any new log messages added to the Samba source files now should appear in the log files.

Determining where DEBUG statements should be placed in the code may require setting various log levels in the smb.conf file. The output of the log files can help narrow down which source files should be examined for particular information. Using printf statements also may help in determining which variables should be logged and in formulating the final log message. If you do plan on using printf statements, smbd and nmbd should be executed without the -D option by stopping the dæmons with the command `/etc/init.d/samba stop` and executing `/usr/local/samba/bin/smbd` and `/usr/local/samba/bin/nmbd` on the command line. The printf statements then are directed toward standard output and appear on the console.

Listings 2 through 4 show custom DEBUG statements added to the Samba source code. Listing 2 shows DEBUG statements added to the source/rpc_server/srv_netlog_nt.c file for reporting failed and successful network logons. The first DEBUG statement reports when an unknown user attempts to log on to the network, and the second DEBUG statement

Listing 3. DEBUG Call Added to source/smbd/service.c

```
// ESK
if(strcmp(lp_servicename(SNUM(conn)), \
                                       "share") &&
   strcmp(lp_servicename(SNUM(conn)), \
                                    "profiles") &&
   strcmp(lp_servicename(SNUM(conn)),\
                                    "netlogon") &&
   strcmp(lp_servicename(SNUM(conn)), "IPC$")){
 DEBUG(0, ("%s (%s) closed connection to \
                                     service %s\n",
      remote_machine,conn->client_address,
      lp_servicename(SNUM(conn))));
  }
```

Listing 4. DEBUG Call Added to source/smbd/chgpasswd.c

```

 /* Logs Password Change */
 if (chstat)
     DEBUG(0, ("Password Change . User:[%s] \
            %sPassword Successfully Changed\n",
          name, (chstat ? "" : "un")));
     return (chstat);
}
```

records incorrect passwords. An additional DEBUG statement was added for installations of Samba using PAM. The final DEBUG statement records a successful logon to the network. By examining the log output from Listing 1, you should see a direct correspondence between each of the DEBUG statements and the generated log entries.

Listing 3 shows a DEBUG statement added to the source/smbd/service.c file to capture when a user has logged off the system by checking when a share to the user has been closed. Unfortunately, this is an unreliable check because the user always is dropping shares during the course of a session. There also is a short delay between the time the user logs off the network and when the dropped share is recorded. Determining when a user has logged off the network requires checking any logons to the machine after the last user share was dropped or checking whether the machine still is locked by the user.

Once Samba is logging the required information, you may want to clean up the log file by removing unnecessary entries. This can be accomplished by setting the log level to 0 in required DEBUG statements and setting the log level to 1 or higher for other DEBUG statements. The log level parameter in the smb.conf file then should be set to 0. The logging features of Samba make it easy to track down unwanted log entries by providing the exact location of the DEBUG statement. Cleaning up log.smbd makes the audits easier and less error-prone than a cluttered log file, such

PORTLOCK
STORAGE MANAGER
ESSENTIAL TOOLS FOR STORAGE

Solutions for Linux and NetWare

Bare-Metal Disaster Recovery
Server-to-Server Migration
Storage Management

A system administrator's time is expensive - downtime is even more expensive. You need products with a proven track record of reliability and functionality, and a company that stands behind those products with outstanding support.

PORTLOCK

www.portlock.com/psm

as the log file generated by a Windows 2000 server.

Updating User Passwords

When updating passwords, system requirements state that passwords must be at least eight characters long and the password change must be logged. In addition, we also wanted the passwords to be synchronized between Windows and Linux so users have common logins for both systems.

For the first requirement, the define statement in source/include/local.h, on line 175, was changed to `#define MINPASSWDLENGTH 8`. To ensure this change is captured in all the necessary source files, `make clean` should be executed in the source directory before executing `make install`.

The source code for verifying and updating password changes is located in the file source/smbd/chgpasswd.c. Listing 4 shows the DEBUG statement that was added to the end of the chat_with_program function to log when users successfully change their passwords. In addition to adding the capability to record successful password changes, the failed password updates also are logged. Failed password changes are recorded because regardless of why the password update failed, the following message always is returned to the user:

```
The User name or old password is incorrect.
Letters in passwords must be typed using the
correct case. Make sure the Caps Lock is not
accidentally on.
```

These log messages can help frustrated users determine why they are unable to update their passwords. However, access to these log messages requires assistance from the system administrator. Log entries 5 and 6 in Listing 1 present two examples of user bill being unable to change his password successfully.

To synchronize passwords between Samba and the Linux system passwords, set the following fields in the smb.conf file under the global section:

```
[global]
unix password sync = yes
pam password change = yes
passwd program = /usr/bin/passwd
passwd chat =*New*password* %n\n *new*password*
➥%n\n *successfully*
```

For most systems, the `passwd chat` field does not need to be set, because the default setting works fine. If the passwd chat field does need to be set, the syntax should follow the passwd command's input and output closely. The syntax for password chat is * for any character and %n for the new password; spaces designate new lines, and \n is used when user input is required. For further help with debugging, set the log level to 101 and the field `passwd chat debug` to yes in the global section of the smb.conf file. As a last resort, printf and DEBUG statements can be used in the function chat_with_program in the chpasswd.c file to help debug the problem.

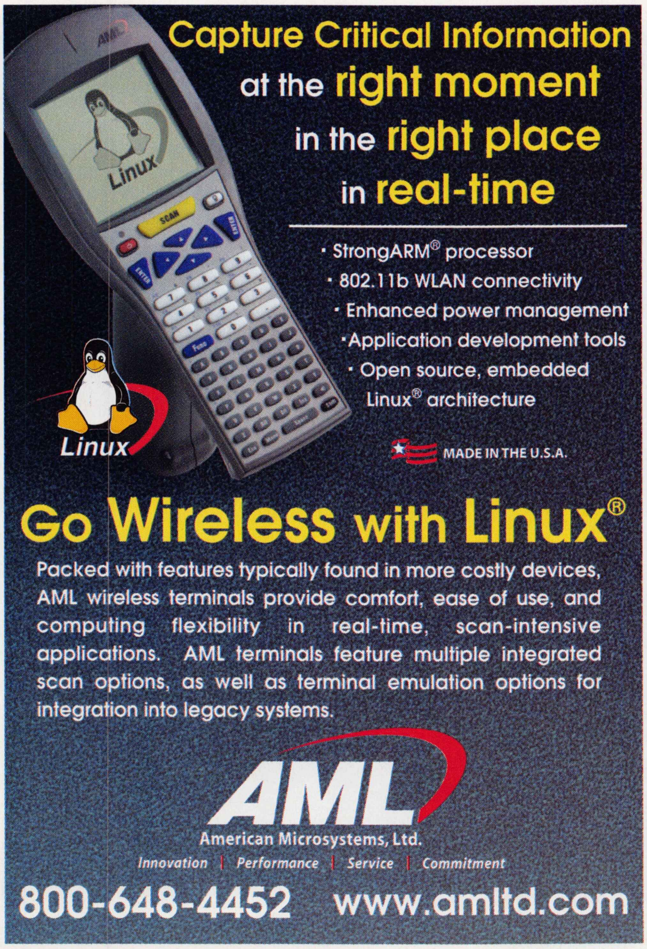
Future Work

Some of the current problems mentioned above, such as displaying a more meaningful error message when users fail to update their passwords and a more exact method of determining when users have logged off the network, need to be addressed. Additional features, such as user lockout after five consecutive unsuccessful login attempts and preventing the user from reusing the five previous passwords, also should be added. Using LDAP for both Linux and Windows clients is worth investigating as well.

Conclusion

Samba provides a reasonable alternative to using Windows 2000 servers on a network to manage Windows clients. The primary advantage of Samba over Windows is the ability to modify the Samba source code to create a system tailored for a specific computing need. It also offers network administrators the ability to troubleshoot the network at the source code level. None of this is possible when using proprietary software, a huge drawback in configuring and debugging services on a network. All of these factors add up to large cost savings in both licensing fees and network administration time.

Edward Kablaoui (eskablaoui@yahoo.com) currently is a software engineer at Northrop Grumman. He lives in Maryland with his wife, Nancy.



Capture Critical Information at the right moment in the right place in real-time

- StrongARM® processor
- 802.11b WLAN connectivity
- Enhanced power management
- Application development tools
- Open source, embedded Linux® architecture

MADE IN THE U.S.A.

Go Wireless with Linux®

Packed with features typically found in more costly devices, AML wireless terminals provide comfort, ease of use, and computing flexibility in real-time, scan-intensive applications. AML terminals feature multiple integrated scan options, as well as terminal emulation options for integration into legacy systems.

AML
American Microsystems, Ltd.
Innovation | Performance | Service | Commitment
800-648-4452 www.amltd.com

NetOp Remote Control

Real Security.
Real Remote Control.

Award-winning NetOp Remote Control lets you view a remote PC's screen, control its keyboard and mouse, synchronize files, launch applications or chat with someone at the remote PC - just as if you were seated at that computer. NetOp -- Secure, cross-platform, support for all your Linux, Solaris & Windows PCs and servers.

Take NetOp for a FREE test flight today

- Remote control a current Linux desktop session
- Fast file transfers
- Advanced security

Get your free evaluation:
CrossTec Corporation
www.CrossTecCorp.com
1-800-675-0729

NetOp® Remote Control

Linux
Apache
MySQL
PHP/Perl

$99/mo

Dedicated Servers
10Mbps Included!

cari.net™
We Manage, You Profit!

Enlightened Hosting Services
USA 1.888.221.5902
www.cari.net/lamp

L3 Systems

KeyRF
RF Remote Control for Presentations

- **RF Technology** - Range to 100 ft at 360°- no pointing
- **Complete Kit** - Transmitter, Receiver and cable included
- **Simple Installation** - Requires no special software or drivers, plugs into PS/2 or USB port and emulates keyboard.
- **Easy to Use** - Two (or four) simple buttons, buttons glow in the dark, & fits easily in a pocket

L3 Systems
PO Box 2954
Redmond WA 98073
Tel: (425) 836-5438

http://www.L3sys.com
email: sales@L3sys.com
Fax: (425) 868-8706

www.ITProSuccess.com

FREE MARKETING ADVICE

LINUX SIMULATOR™
Interactive Simulations for Linux applications

- 100% Programming Free
- Capture using a continuous 'pixel-based' methodology
- Multimedia edit suite including drag-and-drop edit objects
- Generate interactive Flash Files; no proprietary plug-in required

XStream™ SOFTWARE www.xstreamsoftware.com

PENGUIN CAFFEINATED PEPPERMINTS

Who says penguins can't fly?

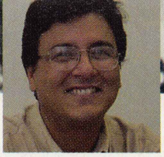
http://store.linuxjournal.com

Linux Media Labs

LMLM4 LMLH264
MPEG-4 Hardware Encoder
H264 and Multichannel
RTSP/RTP/RTCP live streaming
1.0 Mbit/sec for full rate video
720x480 (576) @ 30 (25) fps
1,2,4 channels, MPEG-4 Part10 options

LMLBT44/4M
Multichannel Video IO
4/8 Channels Video Capture : 720x480
4 Input Sensors/ 3 Output Alarms
Fully Supported by Linux Kernel

LINUX MEDIA LABS
www.linuxmedialabs.com
e-mail : sales@linuxmedialabs.com
1-888-795-4735 or 1-719-231-3173
Free Gift with any purchase : Promotion Code #200305LJ



SOLIS, a Brazilian Free Software Cooperative

If a co-op business structure works for Sunkist and Land O'Lakes, it could work for software too. After several successful university software development projects, developers in Brazil are using an old business plan in a new way. **BY CESAR BROD**

Univates, a university center in the south of Brazil, has been a free software user since 1997 and has adopted free software as a standard for administrative people since 1999. As there was no academic administration system available as free software, Univates developed its own, SAGU. SAGU now is used by a dozen other universities in Brazil, with some of them contributing code. By adopting free software and developing SAGU, Univates has saved almost $200,000 US in software licenses. It further saves $70,000 US every year as it doesn't need to buy new licenses or upgrades when installing new computers or increasing the user base. The number of students in Univates has grown from a little less than 2,000 in 1999 to more than 7,500 in 2003. The savings always have been more than enough to cover the payroll of the whole IT department, now employing 26 people.

If you search Google with the keywords Free Software Brazil, you surely will find Univates' projects among the first search results. These projects include a library automation and integrated circulation system, GNUTECA, which is compliant with international library standards. It has been promoted in South America through workshops held by Unesco's office in Montevideo. There also is a PHP-based, object-oriented development framework, MIOLO, that was used to develop GNUTECA and is the base of several other projects. You also should look at Agata Report, a full-featured professional database reporting tool; see www.univates.br/freesoftware.

Regional Development

Univates is a community-owned university. Its only financial support comes from students' payments and some services it provides to the region's industries. It serves a region made up of 40 cities where 300,000 people live. As it is owned by the community, Univates also is audited by this community, which participates in the election of the rector and in several committees that help decide the university's actions. Univates is committed to regional development and has promoted several programs through which it helps local industry.

Univates has watched its IT department become self-sustainable and even generate resources for the university by selling services such as software customization and technology transfer seminars. However, as Univates never wanted to be a software house, it proposed to its IT team that it become a separate entity. As its own entity, it can be hired by Univates and by others, expanding its business and creating job positions to be filled by the university's graduates and students. Both this new entity and Univates would work on a business plan that would link regional development and free software.

The International Co-operative Association (www.ica.coop) says that:

> Co-ops are based on helping each other and caring for others. A co-operative is a type of business or organization. It is a group of people who are working together to solve their own problems and meet their needs. Co-ops are different from other types of organizations since they abide by three main rules: 1) co-ops treat people fairly and respectfully; 2) co-ops encourage people to work together toward solving their mutual problems; and 3) co-ops provide products and services to meet people's needs rather than solely for the purpose of making money.

By this definition, it is quite easy to see that a cooperative organization and the free software philosophy have a lot in common. Both are concerned with treating people with respect and working toward common objectives. The new organization that was born from Univates' IT department is a cooperative called SOLIS (www.solis.coop.br), officially inaugurated on May 12, 2003.

SOLIS works toward regional development by finding ways of empowering local businesses through the use of free software. By doing this, people at SOLIS hope these businesses can become more competitive, profitable and able to grow and hire more people.

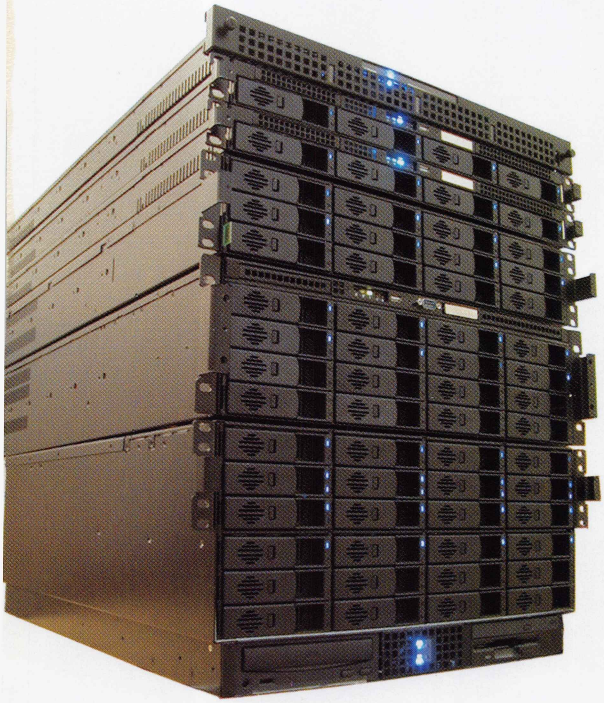
Right now SOLIS is working with industrial and commercial associations, helping the companies associated with them fine-tune their needs in order to save money by sharing among them the costs of software development, customization and training. Univates and SOLIS wish to prove in a practical way that free software can help regional development and is good business that—most importantly—respects people's free access to knowledge. We hope other universities can set up similar arrangements in their own regions.

Cesar Brod (cesar@brod.com.br) first became involved with the GNU/Linux operating system in 1993. He now is the IT manager and coordinates software development for Univates. Cesar also is the Vice President of SOLIS and the coordinator of the international section of the Latin America Free Software Conference (www.softwarelivre.pti.org.br), which will be in November 2004, in the beautiful setting of the Iguassa Falls, among Paraguay, Argentina and Brazil.

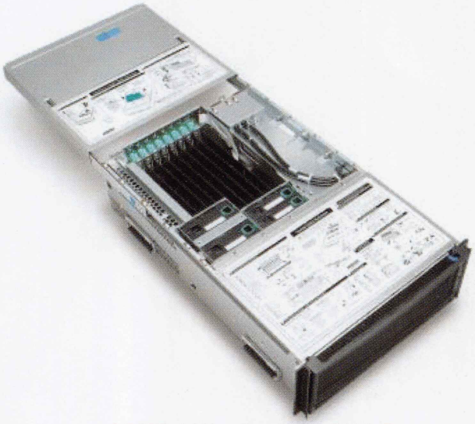
Polywell Servers - Made in USA

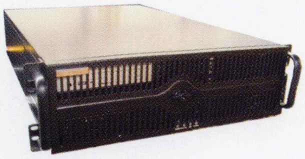
More Choices To Build Your Applications

1U P4 LinuxServer starts **$699**

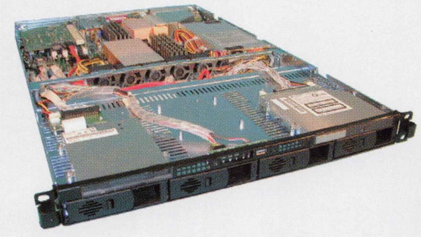

1U to 5U Mass Storage Servers w/up to 24 Drives per Rack

1U 15" LCD +K/B+TrackBall


A6400 4-Way Itanium2 64bit Server

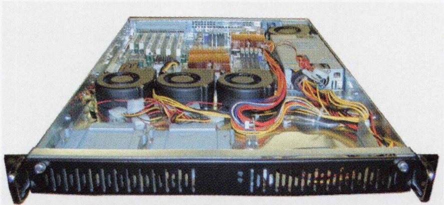

3620BW 3U 6x5.25"+2x3.5"

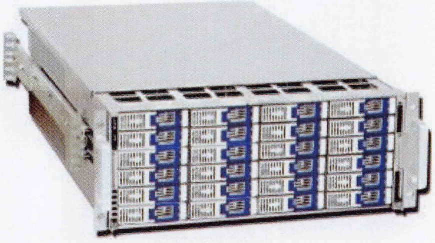
4U 8-Way Xeon MP Server


1114AI 1U 4xSwap+Slim CD/FD

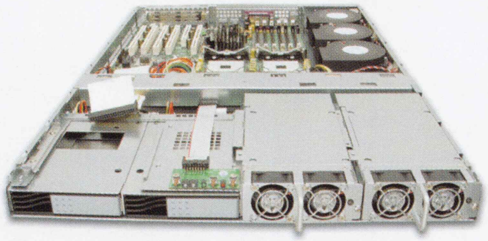
4U 20-Way Xeon Blade Server

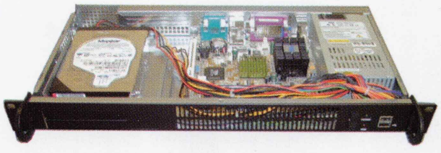
1112PI2 1U 2xPCI Dual CPU Server


1112EV 1U DualOpteron Cluster Node

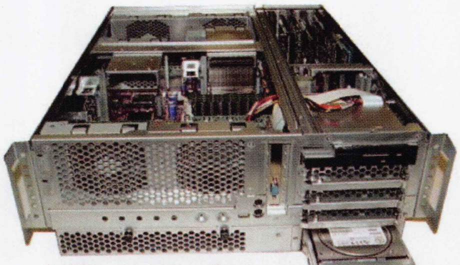

4024AIS 4U 24 Drives NAS Server

1112UB Short 1U with Front USB Ports


1102RP 1U 2xRedundant P/S Dual CPU


1U8i Small 1U 8" Depth Appliance System

Suma 3U 16 Drives Fibre/SCSI/SATA RAID


8400A 4U 4-Way Opteron

15 Years Of Customer Satisfaction
5-Year Industry's Longest Warranty
First Class Customer Services

800.765-9686
www.Polywell.com/us/isp

POLY SYSTEMS

Polywell Computers, Inc 1461 San Mateo Avenue South San Francisco CA 94080 650.583.7222 Fax: 650.583.1974

AMD and OPTERON are trademarks of Advanced Micro Devices, Inc.. Xeon and Itanium is trademark of Intel Corporation. All other brands names are trademarks of their respective companies.

How good is your MPI?

The answer may surprise you.

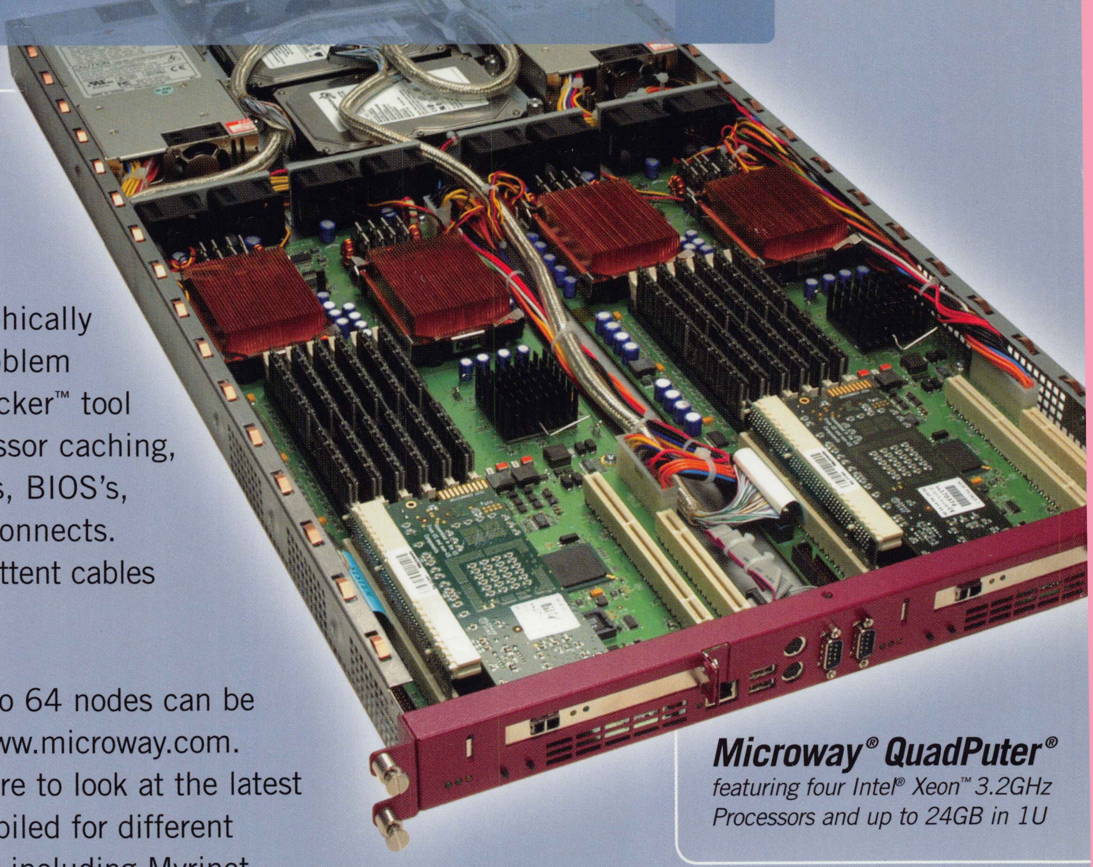
The performance of an MPI application is often limited by the speed of its slowest node! Microway has developed MPI diagnostic tools to help you find the weak link in your cluster.

Cool

High Density

Fast

Reliable



MPI Link-Checker™
measures the bandwidth and latency between all nodes in a cluster then summarizes the data graphically so you can easily spot problem nodes. The MPI Link-Checker™ tool detects issues with processor caching, motherboards, PCI busses, BIOS's, riser cards and PCI interconnects. It can even detect intermittent cables and cross bar switches!

MPI Link-Checker™ for up to 64 nodes can be downloaded gratis from www.microway.com. When you are there, be sure to look at the latest benchmarks we have compiled for different interconnect technologies, including Myrinet, InfiniBand and Gigabit Ethernet and StarGen.

Microway® QuadPuter®
featuring four Intel® Xeon™ 3.2GHz Processors and up to 24GB in 1U

High Density Computing

Microway's XBlade™ powered QuadPuter® offers many unique features for building very dense clusters that will take a full load of DIMM modules, while interfacing to other nodes at speeds approaching a Gigabyte per second. XBlade employs a Microway-ported LinuxBIOS. It is a great building block for creating appliances and Microway has a long history of customizing hardware for COTS and OEM applications.

Excellence in CLUSTER Technology

Please visit Microway.com and discover the very latest in parallel processing, blades and connectivity.

Microway Cluster features:

Nodewatch™
Remote monitoring and control of temperatures, fans and power, including the ability to remotely cycle and reset power via Microway's web-based GUI.

MCMS™
Microway Cluster Management Software provides OS performance data and performs common tasks like adding users, executing cluster-wide commands, and replicating hard disks.

Advanced CoolRack™ Chassis and Cabinets
Robust designs include Processor Targeted Cooling to enable 64 and 32 bit processor based clusters with power densities up to 25 KW per 44U cabinet!

Excellent Technical Support
Since 1982 Microway's Tech Team has been praised by clients for providing accurate and timely pre- and post-sale support.

intel inside ITANIUM 2

intel inside XEON

AMD Opteron

Call us first at 508-746-7341 for quotations and benchmarking services.
Find technical information, testimonials, and online newsletter at www.microway.com

Microway
Technology you can count on™